確　率　論

伊藤　清　著

岩波基礎数学選書

岩　波　書　店

刊行にあたって

——数学的現象の把握を——

現代の数学は形式主義の影響を強くうけていて，数学の本にも，定義，公理，定理，証明を羅列した形式で書かれたものが多い．形式主義によれば，数学はそれ自身は意味をもたない記号を与えられたルールに従って並べて行くゲームに過ぎないということであるが，これは数学の最も本質的なものを見落しているのではなかろうか？　たとえば，公理的構成の規範となったHilbertの幾何学基礎論では，'点'，'直線'，等は意味のない無定義語，すなわち記号であって，'猫'，'雀'，等で置き換えても一向差支えないことになっているが，これは事実に反する．実際，Hadamardが指摘したように，幾何学基礎論には殆んど毎頁に図が載っている．図を描かず，頭の中で図を想像することもせずに，論理だけによってHilbertの幾何学基礎論を理解することは不可能であろう．このように 意味を考えない形式主義の立場に立つ限り，幾何学基礎論さえも理解できないのである．

私の見る所では，数学は実在する数学的現象を記述しているのであって，数学を理解するということは，窮極において，その記述する数学的現象のイメージを言わば感覚的に把握し，形式主義では捕捉できない数学の意味を理解することである．

岩波講座「基礎数学」は現代の数学の基礎部門の基本的な内容を上述の意味で感覚的にわかり易く解説することを目標として編集された．このたびのシリーズはこの講座の中から学部程度の基礎教科に相当するものを選び「岩波基礎数学選書」として編集しなおしたものである．

この機会に章末問題に対して解答・ヒントを付し，学生の勉学に供するようにした．

1990年6月

小 平 邦 彦

目　　次

第4章 独立確率変数の和

第5章 確率過程

はじめに

空間の点は3実数の組であらわされるから，空間図形の幾何学的性質はすべて実変数に関する式であらわされる．したがって代数や解析の知識があれば，図形の性質は論理的には正しく理解できる．しかし幾何学を建設していくには代数や解析の知識だけでは不十分で，図形を直観的に把握しなければならない．確率論の場合も同様である．現代の確率論は，測度論の言葉で叙述することによって論理的には完全に解析学の一分野となっている．しかし確率論を真に楽しむためには，確率現象に対する直観的理解を背景として確率論の発展方向を見きわめなければならない．

本書の第1章では最も簡単な有限試行に限定して，確率現象を解析的に表現する方法を説明し，以下の諸章で一般確率論を展開することにする．

測度論についてはすでに良書も多数出ているし，本選書の中にもその項目があるが，後章で引用するために確率測度の例や諸性質を第2章にまとめておく．

第3章では一般確率論の基礎概念を測度論の言葉で説明する．この考え方は特別の場合についてはすでに第1章でのべてあるから，一般論への移行は，第2章の確率測度の知識があれば，極めて容易である．

第4章以下では，現代確率論の中心問題である確率過程について述べる．限られた紙数でこの理論全体にわたることはできないが，少なくとも基本的な事柄については詳しく述べたつもりである．

注　意

1. 集合の記号．

$\boldsymbol{R}$ または $\boldsymbol{R}^1$ は実数全体，

$\boldsymbol{Q}$ は有理数全体，

$\boldsymbol{N}$ は自然数全体，

集合の和は $\cup$ であらわすが，特に直和は $+$ または $\sum$ であらわす，

集合の差は $\setminus$ であらわすが，特に固有差 ($A \supset B$ のときの $A \setminus B$) は $-$ (すなわち $A - B$) であらわす．

2. “測度 μ に関してほとんど到る所で

$$f_n \longrightarrow f, \quad |f_n| \leqq g \quad \left(\int_S g d\mu < \infty\right)$$

ならば，

$$\int_S f d\mu = \lim_{n\to\infty} \int_S f_n d\mu."$$

この定理は dominated convergence theorem とよばれるが，適当な訳語がないので，有界収束定理と訳すことにした．この訳語は bounded convergence theorem (上の定理で $\mu(S)<\infty$, g=定数 の場合) の意味に用いられるが，この特別の場合をとりたてて用いる必要はおこらない．

3. 例題のヒントは多くの場合略解になっている．いろいろほかに解き方もあろうが，ヒントに従って解くことによりその節で述べた諸定理の応用の練習となると思う．

4. 本書では，位相空間というとき，Hausdorff 位相空間の意味にとる．すなわち，任意の2点 $x, y\,(x \neq y)$ に対し互に素な近傍 $U(x), U(y)$ がとれる(したがってコンパクトな部分空間は常に閉集合である)と仮定する．

第1章 有限試行

有限試行とは，有限個のことなる結果しかもたない試行のことである．本章では対象を有限試行に限定して，確率論の初歩概念を説明する．有限試行を調べるには，有限集合と有限和式を知っておれば十分である．しかも有限試行によって確率論の考え方を理解したならば，解析学の手法を用いて，現代確率論の対象である無限試行に進むのは容易である．

§1.1 確率空間

さいをふってでる目を調べるという試行では，結果は 1, 2, 3, 4, 5, 6 の 6 通りある．このとき，1, 2, …, 6 を**見本点**といい，見本点の全体の集合 $\{1, 2, 3, 4, 5, 6\}$ を**見本空間**という．見本点，見本空間は任意の試行についても考えられる．ある試行の見本空間が有限集合であるか，無限集合であるかに従って，その試行を**有限試行**または**無限試行**という．本章では有限試行だけを考え，これを単に試行という．

Ω をある試行 T の見本空間とする．A を Ω の任意の部分集合とするとき，T の結果として “A に属する見本点が出現する” ことを簡単に A **がおこる**という．このような意味で Ω の部分集合 A を事象 A とよぶこともある．

A の補集合 A^c がおこることと A がおこらないこととは同等である．この意味で A^c を A の**余事象**という．A, B の和集合 $A \cup B$ がおこることと A, B の中少なくとも一つがおこることとは同等であるから，$A \cup B$ を A, B の**和事象**という．A, B の交集合 $A \cap B$ がおこることと A, B が両方ともおこることとは同等であるから，$A \cap B$ を A, B の**交事象**という．A, B の差 $A \setminus B$ がおこることは A がおこってかつ B がおこらないことと同等である．$A \setminus B$ を A, B の**差事象**という．

“$A \subset B$” は “事象 A がおこれば必ず事象 B がおこる” ことを意味する．このときには差 $B \setminus A$ を**固有差**といい，$B-A$ とかく．以後 $B-A$ とかいたときには，$A \subset B$ が暗黙のうちに仮定されているものとする．$A \cap B = \phi$ すなわち A, B が互

いに素であることと事象 A, B が同時にはおこらない(排反事象)こととは同等である．このとき和 $A\cup B$ を直和といい，$A+B$ であらわす．固有差のときと同様に，$A+B$ とかいたときには，A, B が互いに素である(すなわち A, B が排反事象である)ことが暗黙のうちに仮定されているものとする．二つ以上の集合の直和

$$A_1+A_2+\cdots+A_n \qquad \left(\text{または} \sum_{i=1}^{n} A_i \text{ ともかく}\right)$$

についても同様である．

さいをふったときには，見本空間 Ω が

$$\Omega=\{1, 2, 3, 4, 5, 6\}$$

で与えられることは上にのべたが，同時にまた事象 A のおこる確率 $P(A)$ が

$$P(A)=\frac{\#A}{6} \qquad (\#A=A \text{ の点の数})$$

で与えられる．

$\boldsymbol{T}$ を一般の試行とし，Ω をその見本空間とする．任意の $A\subset\Omega$ に対して，事象 A のおこる確率を $P(A)$ とかく．$P(A)$ は集合 A の関数である．この集合関数 P を試行 $\boldsymbol{T}$ の確率法則という．確率の直観的意味から，当然 $P(A)$ は次の性質をもつ．

(P. 1) $\qquad P(A)\geqq 0$,

(P. 2) (加法性) $\qquad P(A+B)=P(A)+P(B)$,

(P. 3) $\qquad P(\Omega)=1$.

一般に有限集合 Ω の上の集合関数 P でこの性質をもつものを Ω の上の確率測度といい，Ω に P をそえたものを確率空間 (Ω, P) という．

上の考察から，試行 $\boldsymbol{T}$ の確率法則 P は $\boldsymbol{T}$ の見本空間 Ω の上の確率測度である．$\boldsymbol{T}$ の見本空間 Ω に $\boldsymbol{T}$ の確率法則 P をそえて得られる確率空間 (Ω, P) を試行 $\boldsymbol{T}$ の確率空間という．$\boldsymbol{T}$ の確率空間 (Ω, P) は $\boldsymbol{T}$ の数学的側面を完全に表わしているから，$\boldsymbol{T}$ を研究するには (Ω, P) を調べたらよい．

定理 1.1 P を Ω の上の確率測度とすれば，

(i) $P\left(\sum_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)$,

(ii) $P(B-A)=P(B)-P(A)$,

(iii) $P(A^c)=1-P(A)$,

(iv) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$,

(v) $P(A)=\sum_{\omega\in A} P\{\omega\}$ (したがってすべての $\omega\in\Omega$ に対して $P\{\omega\}$ を与えると，P が完全に定まる)．

証明 (i) (P. 2) を考慮にいれて帰納法を用いよ．

(ii) $B\supset A$ は暗黙の中に仮定されているから

$$B=(B-A)+A.$$

これに (P. 2) を適用せよ．

(iii) (ii) において $B=\Omega$ とおけ．

(iv) $C=A\cap B,\ A_1=A-C,\ B_1=B-C$ とおくと

$$A=A_1+C, \quad B=B_1+C, \quad A\cup B=A_1+B_1+C.$$

これに (i) を適用せよ．

(v) $A=\sum_{a\in A}\{a\}$ に (i) を適用せよ．▌

事象は集合であらわされるといったが，ときには見本点 ω に関する条件 $\alpha(\omega)$ であらわされることがある．条件 $\alpha(\omega)$ を成立させるような見本点 ω が，試行の結果として出現することを α が**おこる**という．この意味で条件 $\alpha(\omega)$ も事象という．

$A=\{\omega\,|\,\alpha(\omega)\}$ とおくと，α がおこることと A がおこることとは同等である．したがって α のおこる確率は $P\{\omega\,|\,\alpha(\omega)\}$ に等しい．

α の否定条件を $\alpha^{\urcorner}$ であらわす．$\alpha^{\urcorner}$ は α の余事象である．'α または β' という条件を $\alpha\vee\beta$ であらわす．

$$\{\omega\,|\,\alpha(\omega)\vee\beta(\omega)\}=\{\omega\,|\,\alpha(\omega)\}\cup\{\omega\,|\,\beta(\omega)\}$$

であるから，$\alpha\vee\beta$ は α, β の和事象である．同様に 'α かつ β' という条件を $\alpha\wedge\beta$ であらわす．$\alpha\wedge\beta$ は α, β の交事象である．

例題 1.1 確率測度 P に対して次の包含排除公式を証明せよ．

(i) $P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{k=1}^{n}(-1)^{k-1}\sum_{i_1<i_2<\cdots<i_k} P\left(\bigcap_{\nu=1}^{k} A_{i_\nu}\right)$,

(ii) 上の式で $\bigcup$ と $\bigcap$ とを交換して得られる式．

[ヒント] 定理 1.1 の (iv) を考慮に入れて，n に関して帰納法を用いよ．

§1.2　実確率変数，確率ベクトル

(Ω, P) をある試行 T の確率空間とする．Ω の上で定義された実数値関数 $X(\omega)$ を (Ω, P) の上の**実確率変数**という．これは T の結果見本点 ω が出現したとき，$X(\omega)$ という値をとる偶然量をあらわしている．

前節で述べたように‘さいをふる’という試行の確率空間 (Ω, P) は

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \qquad P(A) = \frac{\# A}{6}$$

で与えられる．今さいをふってでた目の2倍の賞金を貰うという賭けをしたとき，この賞金 X は

$$X(\omega) = 2\omega$$

で定義される実確率変数である．

さいを2回つづけてふるという試行の確率空間 (Ω, P) は

$$\Omega = \{(i, j) \mid i, j = 1, 2, 3, 4, 5, 6\} = \{1, 2, 3, 4, 5, 6\}^2,$$

$$P(A) = \frac{\# A}{36}$$

で与えられる．この試行で1回目にあらわれる目を X_1，2回目にあらわれる目を X_2，その和を X とすると，これらは

$$X_1(i, j) = i, \qquad X_2(i, j) = j, \qquad X(i, j) = i + j$$

で定義される実確率変数である．

さてもとの一般の話にもどり，$X = X(\omega)$ を (Ω, P) の上の実確率変数とする．$X(\omega)$ のとり得るすべての値の集合すなわち $X(\Omega)$ を X の**見本空間**といい，Ω^X であらわす．任意の $B \subset \Omega^X$ に対して，“X の値が B に入る”という事象の確率は $P\{\omega \mid X(\omega) \in B\}$ すなわち $P(X^{-1}(B))$ に等しい．これを B の関数とみて $P^X(B)$ とかくと，P^X は Ω^X の上の集合関数として確率測度の条件をみたしていることは

$$X^{-1}(B_1 + B_2) = X^{-1}(B_1) + X^{-1}(B_2), \qquad X^{-1}(\Omega^X) = \Omega$$

から容易にわかる．確率測度 P^X を X の**確率法則**という．とくに

$$P^X\{b\} = P(X^{-1}(b)), \qquad b \in \Omega^X.$$

P^X は Ω^X の上の確率測度であるから，任意の $B \subset \Omega^X$ に対して

$$P^X(B) = \sum_{b \in B} P^X\{b\}$$

となり，P^X は $P^X\{b\}$ $(b \in \Omega^X)$ を与えることによって完全にきまる．

上のさいを2回ふる例であらわれた実確率変数 X_1, X_2, X の確率空間を求めて見よう．図1.1からわかるように

$$\Omega^{X_1} = \Omega^{X_2} = \{1, 2, 3, 4, 5, 6\}, \qquad \Omega^X = \{2, 3, 4, \cdots, 12\},$$

$$P^{X_i}\{k\} = P(X_i^{-1}(k)) = \frac{\# X_i^{-1}(k)}{36} = \frac{6}{36} = \frac{1}{6}, \qquad i = 1, 2,$$

$$P^X\{k\} = P(X^{-1}(k)) = \frac{\# X^{-1}(k)}{36} = \begin{cases} (k-1)/36 & (2 \leqq k \leqq 7), \\ (13-k)/36 & (8 \leqq k \leqq 12). \end{cases}$$

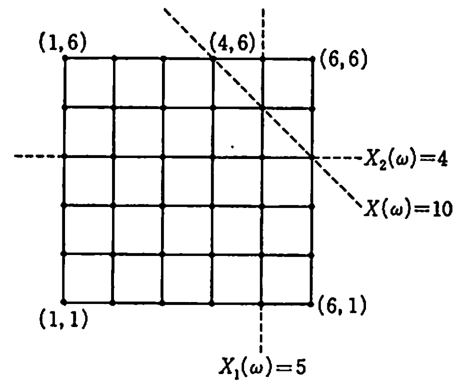


図1.1

もとの一般の確率空間 (Ω, P) にもどる．実確率変数 $X_1(\omega), X_2(\omega)$ をならべると，2次元の値をとる関数

$$X(\omega) = (X_1(\omega), X_2(\omega))$$

が得られる．これを2次元の**確率ベクトル変数**（または簡単に**確率ベクトル**）という．このときにも X の見本空間 Ω^X，確率法則 P^X を

$$\Omega^X = X(\Omega),$$

$$P^X(B) = P\{\omega \mid X(\omega) \in B\} = P(X^{-1}(B)), \qquad B \subset \Omega^X$$

で定義する．Ω^X は2次元の空間 R^2 の部分集合で，P^X は Ω^X の上の確率測度である．

$$\Omega^X \subset \Omega^{X_1} \times \Omega^{X_2}$$

であるが, 必ずしもこの両辺の集合は一致しない. $X_1(\omega), X_2(\omega)$ を $X(\omega)$ の成分変数といい, $X(\omega)$ を $X_1(\omega), X_2(\omega)$ の**結合変数**という. $X(\omega)$ の確率法則 P^X を X_1, X_2 の**結合確率法則**(または**結合分布**)という. 一般の自然数 n に対して, n **次元の確率ベクトル**も同様に定義する.

n 次元の空間 $\boldsymbol{R}^n$ の点 $(x_1, x_2, \cdots, x_n)$ にその第 k 成分 x_k を対応させる関数(写像といってもよい)を k **射影**といい, π_k であらわす. すなわち

$$\pi_k(x_1, x_2, \cdots, x_n) = x_k.$$

$X(\omega)$ が $X_1(\omega), X_2(\omega), \cdots, X_k(\omega)$ の結合変数であれば,

$$X_k(\omega) = \pi_k(X(\omega)) \quad \text{したがって} \quad X_k = \pi_k \circ X.$$

ここで小丸 $\circ$ は写像の結合を示す.

$X(\omega)$ を確率変数とし, Ω^X, P^X をその見本空間, 確率法則という. Ω^X の上に定義された実数値関数 $\varphi : \Omega^X \to \boldsymbol{R}^1$ に対して,

$$Y(\omega) = \varphi(X(\omega)) \quad \text{すなわち} \quad Y = \varphi \circ X$$

とすると, 新しい実確率変数 $Y(\omega)$ が得られる. Y の見本空間 Ω^Y, 確率法則 P^Y はもちろん

$$\Omega^Y = (\varphi \circ X)(\Omega),$$
$$P^Y(C) = P((\varphi \circ X)^{-1}(C))$$

で与えられる. このようにして確率変数 $X(\omega)$ から得られる実確率変数 $Y(\omega)$ を $X(\omega)$ の関数であるという. もっと一般に $X(\omega)$ が n 次元の確率ベクトル, φ が $\Omega^X(\subset \boldsymbol{R}^n)$ から $\boldsymbol{R}^m$ の中への写像のときには, $Y(\omega) = \varphi(X(\omega))$ は m 次元の確率ベクトルとなる. このときにも $Y(\omega)$ は $X(\omega)$ の関数であるという. とくに φ として k 射影 π_k をとると, $\pi_k(X(\omega))$ は $X(\omega)$ の k 成分 $X_k(\omega)$ に等しく, したがって $X_k(\omega)$ は $X(\omega)$ の関数である.

定理 1.2 $Y(\omega) = \varphi(X(\omega))$ のときには

$$\Omega^Y = \varphi(\Omega^X), \qquad P^Y(C) = P^X(\varphi^{-1}(C)).$$

証明 $Y(\omega) = \varphi(X(\omega)) = (\varphi \circ X)(\omega)$ であるから

$$\Omega^Y = (\varphi \circ X)(\Omega) = \varphi(X(\Omega)) = \varphi(\Omega^X),$$
$$P^Y(C) = P((\varphi \circ X)^{-1}(C)) = P(X^{-1}(\varphi^{-1}(C))) = P^X(\varphi^{-1}(C)).$$

上の推論で $(\varphi \circ X)^{-1}(C) = X^{-1}(\varphi^{-1}(C))$ を用いたが, その証明は

$$\omega \in (\varphi \circ X)^{-1}(C) \iff (\varphi \circ X)(\omega) \in C \iff \varphi(X(\omega)) \in C$$

$$\iff X(\omega) \in \varphi^{-1}(C) \iff \omega \in X^{-1}(\varphi^{-1}(C)). \qquad \blacksquare$$

実確率変数 $X(\omega)$ に対して

$$EX = \sum_{\omega \in \Omega} X(\omega) P\{\omega\}$$

を X の**平均値**(または**期待値**)という. A を Ω の任意の集合とするとき

$$\sum_{\omega \in A} X(\omega) P\{\omega\}$$

を $E(X, A)$ とかく. $X(\omega)$ が確率ベクトルのときに X の**平均値ベクトル** EX を成分ごとに定義する. すなわち $X(\omega)$ の k 成分が $X_k(\omega)$ であれば,

$$EX = (EX_1, EX_2, \cdots, EX_n) \in \boldsymbol{R}^n$$

である.

定理 1.3 X, Y が確率ベクトルならば,

(i) (**平均値の加法性**) $E(aX+bY) = aEX+bEY$ (a, b は定数),

(ii) $E\left(X, \sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} E(X, A_i)$,

(iii) A の上で $X(\omega) \equiv a$ (a は定ベクトル) ならば $E(X, A) = aP(A)$ とくに $E(a) = a$ (左辺の a は恒等的に a に等しい確率ベクトル),

(iv) $EX = \sum_{x \in \Omega^X} x P^X\{x\}$,

(v) $Y(\omega) = \varphi(X(\omega))$ ならば, $EY = \sum_{x \in \Omega^X} \varphi(x) P^X\{x\}$.

X, Y が実確率変数ならば,

(vi) $X(\omega) \geqq Y(\omega) \implies EX \geqq EY$, $E(X, A) \geqq E(Y, A)$,

(vii) $X(\omega) \geqq 0$, $A \subset B \implies E(X, A) \leqq E(X, B)$.

証明 (iv), (v) 以外は容易であるから読者にまかせ, (iv), (v) だけを証明する.

(iv) $A_x = X^{-1}(x)$ とおくと,

$$\Omega = \sum_{x \in \Omega^X} A_x,$$

したがって

$$E(X) = E(X, \Omega) = \sum_{x \in \Omega^X} E(X, A_x).$$

A_x の上では $X(\omega) = x$ であるから,

$$E(X, A_x) = xP(A_x) = xP^X\{x\}.$$

この2式から (iv) の式が得られる.

(v) 上と同様にして

$$E(\varphi(X)) = \sum_{x\in\Omega^X} E(\varphi(X), A_x) = \sum_{x\in\Omega^X}\varphi(x)P(A_x) = \sum_{x\in\Omega^X}\varphi(x)P^X\{x\}$$

で，(v)の式が得られる．∎

A を Ω の任意の部分集合とするとき，A の指示関数 1_A を

$$1_A(\omega) = \begin{cases} 1, & \omega\in A, \\ 0, & \omega\in A^c \end{cases}$$

で定義する．これは事象 A がおこれば1，おこらなければ0という値をとる確率変数である．明らかに

$$1_\Omega(\omega) \equiv 1.$$

定理1.3の(iii)で恒等的に a に等しい確率変数を a であらわしたが，正確にいえば $a1_\Omega(\omega)$ とかくべきものであった．しかしこのような言葉の乱用は数学全般にわたって普通であり，今後も混乱のおそれがない限り，$a1_\Omega(\omega)$ を単に a とかくことにする．明らかに

$$E1_A = P(A), \quad E(X, A) = E(X1_A).$$

定理 1.4 (i) $1_{A^c} = 1-1_A$,

(ii) $A=\bigcap_{i=1}^n A_i$ ならば，

$$1_A = \prod_{i=1}^n 1_{A_i} = \min(1_{A_1}, 1_{A_2}, \cdots, 1_{A_n}),$$

(iii) $A=\bigcup_{i=1}^n A_i$ ならば，

$$1_A = 1-\prod_{i=1}^n(1-1_{A_i}) = \max(1_{A_1}, 1_{A_2}, \cdots, 1_{A_n}).$$

証明 (i), (ii)は容易．(iii)の始めの等式を示すには，(i), (ii)を用いて，

$$1-1_A = 1_{A^c} = \prod_{i=1}^n 1_{A_i^c} \qquad \left(\left(\bigcup_{i=1}^n A_i\right)^c = \bigcap_{i=1}^n A_i^c \text{ に注意せよ}\right)$$

$$= \prod_{i=1}^n (1-1_{A_i})$$

とすればよい．また 1_A が $\max(1_{A_1}, 1_{A_2}, \cdots, 1_{A_n})$ に等しいことは容易にわかる．∎

確率変数 X の分散 $V(X)$ を

$$V(X) = E((X-EX)^2)$$

で定義する．定理1.3の(v)で $\varphi(x)=(x-EX)^2$ とおいて，

$$V(X) = \sum_{x\in\Omega^X}(x-E(X))^2P^X\{x\}.$$

二つの実確率変数 X, Y の共分散 $V(X, Y)$ を

$$V(X, Y) = E((X-EX)(Y-EY))$$

で定義する．定理1.3の(v)で X のかわりに結合変数 (X, Y) をとり，$\varphi(x, y)=(x-EX)(y-EY)$ とおいて，

$$V(X, Y) = \sum_{(x,y)\in\Omega^{(X,Y)}}(x-EX)(y-EY)P^{(X,Y)}\{(x, y)\}.$$

明らかに $V(X, Y)=V(Y, X)$，$V(X, a)=0$ (a は定数)，$V(X, X)=V(X)\geqq 0$.

定理 1.5 (i) $V(aX+b) = a^2V(X)$, $\quad V(aX+b, cY+d) = acV(X, Y)$,

(ii) $V(aX+bY) = a^2V(X)+2abV(X, Y)+b^2V(Y)$,

(iii) $V(X) = EX^2-(EX)^2$, $\quad V(X, Y) = E(XY)-EXEY$,

(iv) $|V(X, Y)| \leqq \sqrt{V(X)V(Y)}$.

証明 (i), (ii), (iii)を示すには次の式の両辺の平均値をとればよい．

$$((aX+b)-E(aX+b))^2 = a^2(X-E(X))^2,$$
$$((aX+b)-E(aX+b))((cY+d)-E(cY+d)) = ac(X-EX)(Y-EY),$$
$$((aX+bY)-E(aX+bY))^2 = a^2(X-EX)^2+2ab(X-EX)(Y-EY)$$
$$+b^2(Y-EY)^2,$$
$$(X-EX)^2 = X^2-2(EX)\cdot X+(EX)^2,$$
$$(X-EX)(Y-EY) = XY-(EX)Y-(EY)X+EXEY.$$

(iv)を示すには，任意の実数 t に対して

$$\begin{aligned} 0 \leqq V(tX+Y) &= E((tX+Y)-E(tX+Y))^2 \\ &= E(t^2(X-EX)^2+2t(X-EX)(Y-EY)+(Y-EY)^2) \\ &= t^2V(X)+2tV(X, Y)+V(Y). \end{aligned}$$

したがって最後の t の2次式は t のいかんにかかわらず非負．判別式をとると，

$$V(X, Y)^2-V(X)V(Y) \leqq 0 \quad \text{すなわち} \quad |V(X, Y)| \leqq \sqrt{V(X)V(Y)}. \ ∎$$

実確率変数 $X(\omega), Y(\omega), Z(\omega)$ に関する条件，例えば

$$X(\omega)+Y(\omega) \geqq Z(\omega)$$

は ω に関する条件でもあり，したがって事象と考えられる．その確率は

$$P\{\omega \mid X(\omega)+Y(\omega) \geqq Z(\omega)\}$$

であるが，これを簡単に

$$P\{X(\omega)+Y(\omega)\geqq Z(\omega)\} \quad \text{または} \quad P\{X+Y\geqq Z\}$$

とかく．

事象 $\alpha(\omega)$ の確率が1であるとき，“ほとんど確実に $\alpha(\omega)$ がなりたつ”といい，記号では

$$\alpha(\omega) \quad \text{a.s.} \qquad (\text{a.s.}=\text{almost surely})$$

とかく．またこれを“$\alpha(\omega)$ が確率1でなりたつ”ともいう．

X の標準偏差 $\sigma(X)$，X, Y の相関係数 $R(X, Y)$ を

$$\sigma(X)=\sqrt{V(X)}, \qquad R(X, Y)=\frac{V(X, Y)}{\sigma(X)\sigma(Y)} \qquad (\sigma(X), \sigma(Y)>0 \text{ のとき})$$

で定義する．

$$X(\omega)=X'(\omega) \text{ a.s.} \Longrightarrow EX=EX',\ V(X)=V(X'),\ \sigma(X)=\sigma(X'),$$

$$X(\omega)=X'(\omega) \text{ a.s.},\ Y(\omega)=Y'(\omega) \text{ a.s.}$$

$$\Longrightarrow V(X, Y)=V(X', Y'),\ R(X, Y)=R(X', Y').$$

これを示すには E についてだけ証明すれば十分である．$A=\{\omega\,|\,X(\omega)\neq X'(\omega)\}$ とおけば，$P(A)=0$，したがって任意の $\omega\in A$ に対して $P\{\omega\}=0$．故に $E(X, A)=E(X', A)=0$，$\omega\in A^c$ に対しては $X(\omega)=X'(\omega)$．故に $E(X, A^c)=E(X', A^c)$．故に $EX=EX'$ がでる．

定理 1.6 $\quad X(\omega)\geqq 0,\ EX=0 \Longrightarrow X(\omega)=0$ a.s.

証明 $A=\{\omega\,|\,X(\omega)>0\}$ とすると，$\omega\in A^c$ の上では $X(\omega)=0$．故に

$$EX=E(X, A)=\sum_{\omega\in A}X(\omega)P\{\omega\}.$$

もし $P(A)>0$ ならば，$P\{\omega_0\}>0$ なる $\omega_0\in A$ がある．もちろん $X(\omega_0)>0$．またすべての ω に対して $X(\omega)\geqq 0$ であるから，$EX\geqq X(\omega_0)P\{\omega_0\}>0$．故に $EX=0$ ならば $P(A)=0$，すなわち $X(\omega)=0$ a.s. ▌

定理 1.7 (i) $\sigma(X)=0 \Leftrightarrow X(\omega)=EX$ a.s.,

(ii) $-1\leqq R(X, Y)\leqq 1$,

(iii) $R(X, Y)=\pm 1$ ならば

$$Y(\omega)-EY=\pm\frac{\sigma(Y)}{\sigma(X)}(X(\omega)-EX) \text{ a.s.} \qquad (\text{複号同順}).$$

証明 (i) $\sigma(X)^2=E(X-EX)^2$ であるから，定理1.6により明らか．

(ii) 定理1.5の(iv)により明らか．

(iii) $k=\sigma(Y)/\sigma(X)$ とおく．定理1.5の(ii)と $\sigma(X), R(X, Y)$ の定義により，

$$\begin{aligned}\sigma(Y-kX)^2&=\sigma(Y)^2-2k\sigma(X)\sigma(Y)R(X, Y)+k^2\sigma(X)^2\\&=\sigma(Y)^2-2\sigma(Y)^2R(X, Y)+\sigma(Y)^2\\&=2\sigma(Y)^2(1-R(X, Y)),\end{aligned}$$

したがって

$$\begin{aligned}R(X, Y)=1 &\Leftrightarrow \sigma(Y-kX)=0 \Leftrightarrow Y(\omega)-kX(\omega)=EY-kEX \text{ a.s.}\\&\Leftrightarrow Y(\omega)-EY=k(X(\omega)-EX) \text{ a.s.}\end{aligned}$$

同様にして $R(X, Y)=-1$ の場合も論ずることができる．▌

定理 1.8(Čebyšev の不等式) $a>0$ に対し，

$$P\{|X(\omega)-EX|>a\sigma(X)\}\leqq\frac{1}{a^2},$$

すなわち

$$P\{EX-a\sigma(X)\leqq X(\omega)\leqq EX+a\sigma(X)\}\geqq 1-\frac{1}{a^2}.$$

証明 $\sigma(X)=0$ のときには，前定理の(i)により，$X(\omega)=EX$ a.s. したがって上の不等式は当然なりたつ．$\sigma(X)>0$ とせよ．

$$A=\{\omega\,|\,|X(\omega)-EX|>a\sigma(X)\}$$

とおくと，

$$\sigma(X)^2=E(X-EX)^2\geqq E((X-EX)^2, A)\geqq a^2\sigma(X)^2P(A),$$

故に $P(A)\leqq 1/a^2$ となる．▌

例題 1.2 (ii)以外では X, Y は実確率変数である．

(i) $|EX|\leqq E|X|$，$|EX|\leqq\sqrt{EX^2}$ を証明せよ．

[ヒント] 始めの式は $-|X|\leqq X\leqq |X|$ からでる，あとの式は $EX^2-(EX)^2=V(X)$ からでる．

(ii) 確率ベクトル $X(\omega)$ に関して，$\|EX\|\leqq E\|X\|$ を示せ．($\|x\|$ はベクトル x の長さ)．

[ヒント] $0\leqq E(\|X-EX\|^2)=E(\|X\|^2)-\|EX\|^2$.

(iii) $E1_A=P(A)$ と定理1.4とを用いて，例題1.1(包含排除公式)の(i)の別証を与えよ．これから，$P(A^c)=1-P(A)$ を用いて，同公式(ii)を導け．この際

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}=(1-1)^n=0$$

を利用せよ．

(iv) $\sigma(X)>0$, $a\neq 0$ のとき $R(X, aX+b)=\mathrm{sgn}\, a$ を証明せよ．ただし $\mathrm{sgn}\, a$ は $a\gtrless 0$ に応じて ± 1 をとるような a の関数である．

(v) Čebyšev の不等式を利用して

$$P\left\{\left|\frac{Y(\omega)-EY}{\sigma(Y)}-\frac{X(\omega)-EX}{\sigma(X)}\right|>a\sqrt{2(1-R(X,Y))}\right\}\leqq\frac{1}{a^2}\qquad(a>0)$$

を証明せよ．ただし $\sigma(X), \sigma(Y)>0$ と仮定する．

(vi) さいを2回ふってでる目を X_1, X_2 とするとき，

$$R(a_1X_1+a_2X_2,\ b_1X_1+b_2X_2)$$

を a_1, a_2, b_1, b_2 であらわせ．

(vii) $f(t)=E(X-t)^2$ を最小ならしめる t の値と $f(t)$ の最小値を求めよ．

(viii) $g(t,s)=E(Y-tX-s)^2$ を最小ならしめる t, s の値と $g(t,s)$ の最小値を求めよ．ただし $\sigma(X)>0$ と仮定する．

§1.3 混合，直結合，樹形結合

a) 一般の値をとる確率変数

前節では実確率変数と確率ベクトルを考えた．同様の考えで一般の値をとる確率変数を考えることができる．例えば，さいをふったときでた目が偶数か奇数かに応じて偶，奇という値をとる確率変数も考えられる．さいをふるという試行の確率空間は前にものべたように

$$\Omega=\{1,2,3,4,5,6\},\qquad P\{i\}\equiv\frac{1}{6}$$

で与えられたが，今考える確率変数 X は

$$X(1)=X(3)=X(5)=\text{奇},\qquad X(2)=X(4)=X(6)=\text{偶}$$

という Ω の上の関数であらわされる．このとき X の見本空間 Ω^X は

$$\Omega^X=\{\text{奇},\ \text{偶}\},$$

確率法則 P^X は

$$P^X\{\text{奇}\}=P\{\omega\,|\,X(\omega)=\text{奇}\}=P\{1,3,5\}=\frac{3}{6}=\frac{1}{2},$$

$$P^X\{\text{偶}\}=P\{\omega\,|\,X(\omega)=\text{偶}\}=P\{2,4,6\}=\frac{3}{6}=\frac{1}{2}$$

で与えられる．

また $1,2,\cdots,10$ の番号札が1枚ずつはいった箱から1枚ぬきだすとする．このときの確率空間 (Ω, P) は

$$\Omega=\{1,2,\cdots,10\},\qquad P\{i\}=\frac{1}{10}$$

で与えられる．さて1,2は赤札，3,4,5は青札，残りの5枚は白札とする．このぬきだしの試行ででた札の色を Y とすると，Y は

$$Y(1)=Y(2)=\text{赤},\qquad Y(3)=Y(4)=Y(5)=\text{青},$$
$$Y(6)=Y(7)=\cdots=Y(10)=\text{白}$$

で与えられる確率変数で，その見本空間 Ω^Y と確率法則 P^Y は

$$\Omega^Y=\{\text{赤},\ \text{青},\ \text{白}\},$$

$$P^Y\{\text{赤}\}=\frac{2}{10}=\frac{1}{5},\qquad P^Y\{\text{青}\}=\frac{3}{10},\qquad P^Y\{\text{白}\}=\frac{5}{10}=\frac{1}{2}$$

で与えられる．

一般の値をとる確率変数についても，その見本空間に確率法則をそえて確率空間を定義することは前節の場合と同様であるが，平均値は定義できない．

b) 試行の混合

T を試行，(Ω, P) をその確率空間，$X(\omega)$ を (Ω, P) の上の一つの確率変数とする．T の結果を $X(\omega)$ の値だけに着目して観察すると，新しい試行が得られる．これを T_X であらわし，T の $X(\omega)$ による**混合**という．さいをふって目の偶奇だけに着目するとか，箱の中から札をぬきだしてその色だけに着目するなど，混合の例はいくらもある．T_X の見本空間はもちろん X の見本空間 Ω^X と一致する．T_X の結果 $B(\subset\Omega^X)$ の点がでるというのはもとの T にもどって考えれば，$X^{-1}(B)$ の点がでることであるから，その確率は $P(X^{-1}(B))$ すなわち $P^X(B)$ である．したがって T_X の確率法則は X の確率法則 P^X と一致する．T_X を考えることは Ω の中で $X^{-1}(x)\,(x\in\Omega^X)$ の点をすべて同一視して 'x' という名で呼ぶことであって，これが混合という名称のおこりである．

(Ω^X, P^X) の上の関数 $\varphi(x)$ は T_X に関して確率変数となっている．その確率法則——$(P^X)^\varphi$ であらわす——は

$$(P^X)^\varphi(\Lambda) = P^X(\varphi(x) \in \Lambda) = P^X(\varphi^{-1}(\Lambda))$$

で与えられる．もし $\varphi(x)$ が実数値（またはベクトル値）のときには，その平均値——$E^X\varphi$ であらわす——は

$$E^X\varphi = \sum_{x \in \Omega^X} \varphi(x) P^X\{x\}$$

で与えられる．同じ確率変数を T に関して考えると，Ω の上の関数 $Y(\omega) = \varphi(X(\omega))$ で与えられる．なぜならば，T の結果 ω がでたときには，T_X の結果 $\varphi(X(\omega))$ がでるからである．このように考えたときには，その確率法則はもちろん

$$P^Y(\Lambda) = P(Y^{-1}(\Lambda))$$

で与えられ，φ が実数値のときには，その平均値はもちろん

$$EY$$

で与えられる．$(P^X)^\varphi$ と P^Y，$E^X\varphi$ と EY とはそれぞれ一致すべきであるが，実際それは次の定理によって保証される．

定理 1.9 $Y(\omega) = \varphi(X(\omega))$ とすると

(i) $P\{Y \in \Lambda\} = P^X\{\varphi \in \Lambda\}$,

(ii) φ が実数値（またはベクトル値）のときには

$$EY = E^X\varphi.$$

証明 (i) $P\{Y \in \Lambda\} = P(Y^{-1}(\Lambda)) = P((\varphi \circ X)^{-1}(\Lambda)) = P(X^{-1}(\varphi^{-1}(\Lambda)))$
$$= P^X(\varphi^{-1}(\Lambda)) = P^X\{\varphi \in \Lambda\}.$$

(ii) 定理 1.3 (v) と同様に

$$EY = \sum_{x \in \Omega^X} \varphi(x) P^X\{x\}$$

が証明できるが，右辺は $E^X\varphi$ である．∎

分散，共分散，標準偏差，相関係数などはいずれも平均値をもとにして定義されるから，これらについても上の定理の (ii) と同様な性質がなりたつ．すなわち

$$Y(\omega) = \varphi(X(\omega)), \qquad Z(\omega) = \psi(X(\omega))$$

ならば

$$V(Y) = V^X(\varphi), \qquad V(Y, Z) = V^X(\varphi, \psi),$$
$$\sigma(Y) = \sigma^X(\varphi), \qquad R(Y, Z) = R^X(\varphi, \psi).$$

c) 試行の直結合

T_1, T_2 を試行とし，それぞれの確率空間を $(\Omega_1, P_1), (\Omega_2, P_2)$ とする．この二つの試行をつづけて試み，それをひとまとめにして得られる大きい試行を $\tilde{T}$ とする．$\tilde{T}$ を T_1, T_2 の直結合といい，$T_1 \times T_2$ であらわす．$\tilde{T}$ の結果 ω は T_1 の結果 ω_1 と T_2 の結果 ω_2 との組であらわされるから，$\tilde{T}$ の見本空間 $\tilde{\Omega}$ は

$$\tilde{\Omega} = \{(\omega_1, \omega_2) \mid \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

である．集合論の言葉でいえば，$\tilde{\Omega}$ は Ω_1, Ω_2 の直積で

$$\Omega_1 \times \Omega_2$$

であらわされる．次に $\tilde{T}$ の確率法則 $\tilde{P}$ を考えよう．$T_i\,(i=1,2)$ は $\tilde{T}$ の結果 $\tilde{\omega} = (\omega_1, \omega_2)$ を第 i 成分

$$\omega_i = \pi_i(\tilde{\omega}) \qquad (\pi_i \text{ は } i \text{ 射影})$$

だけに着目して観察したものである．換言すれば，T_i は $\tilde{T}$ の $\pi_i(\tilde{\omega})$ による混合である．したがって当然

$$\tilde{P}\pi_i^{-1} = P_i, \qquad i = 1, 2$$

をみたすべきである．このような $\tilde{P}$ は一般に沢山ある．例えば，硬貨を2回投げるときには

$$\Omega_i = \{0, 1\} \qquad (0 \text{ は裏，} 1 \text{ は表}),$$

$$P_i\{j\} = \frac{1}{2}, \qquad j = 0, 1,$$

$$\tilde{\Omega} = \{(0,0), (0,1), (1,0), (1,1)\}$$

であるが，上の条件をみたす $\tilde{\Omega}$ の上の確率測度 $\tilde{P}$ は

$$\tilde{P}\{(0,0)\} = \alpha, \qquad \tilde{P}\{(0,1)\} = \frac{1}{2} - \alpha,$$

$$\tilde{P}\{(1,0)\} = \frac{1}{2} - \alpha, \qquad \tilde{P}\{(1,1)\} = \alpha$$

$$(\alpha \text{ は } [0, 1/2] \text{ の中の任意の数})$$

で与えられ，無限に多く存在する．

さて一般の場合にもどって $\tilde{P}$ を定めるには，$\tilde{P}\pi_i^{-1} = P_i$ をみたすような $\tilde{P}$ の

中から一つを選ばなければならない．その方法として

乗法律 $\tilde{P}\{(\omega_1, \omega_2)\} = P_1\{\omega_1\}P_2\{\omega_2\}$

を用いる．したがって，任意の $\tilde{A}\subset\tilde{\Omega}$ に対して

$$\tilde{P}(\tilde{A}) = \sum_{\tilde{\omega}\in\tilde{A}} \tilde{P}\{\tilde{\omega}_1\} = \sum_{(\omega_1,\omega_2)\in A} P_1\{\omega_1\}P_2\{\omega_2\}$$

で定義する．こうすると

$$\tilde{P}(A_1\times A_2) = P_1(A_1)P_2(A_2)$$

となり，特に

$$\tilde{P}(\tilde{\Omega}) = \tilde{P}(\Omega_1\times\Omega_2) = P_1(\Omega_1)P_2(\Omega_2) = 1,$$
$$\tilde{P}(\pi_1^{-1}(A_1)) = \tilde{P}(A_1\times\Omega_2) = P_1(A_1)P_2(\Omega_2) = P_1(A_1),$$
$$\tilde{P}(\pi_2^{-1}(A_2)) = P_2(A_2)$$

となって，確かに $\tilde{P}$ は $\tilde{P}\pi_i^{-1}=P_i\,(i=1, 2)$ をみたす $\tilde{\Omega}$ 上の確率測度となる．上の硬貨の例では，この $\tilde{P}$ は $\alpha=1/4$ のときにあたる．

ここでなぜ乗法律を用いたのかという疑問がおこるが，これは直結合に関する**公理**として認めるべきもので，証明できることではない．これは§1.1で確率 P に対して (P. 1), (P. 2), (P. 3) を要請したのと同じである．

一般に乗法律で定義された $\tilde{P}$ を P_1, P_2 の**直積**といい，$P_1\times P_2$ であらわす．この記号を用いると，直結合 $\tilde{\boldsymbol{T}}=\boldsymbol{T}_1\times\boldsymbol{T}_2$ の確率空間 $(\tilde{\Omega}, \tilde{P})$ は

$$\tilde{\Omega} = \Omega_1\times\Omega_2, \qquad \tilde{P} = P_1\times P_2$$

で与えられる．またこの $(\tilde{\Omega}, \tilde{P})$ を (Ω_1, P_1), (Ω_2, P_2) の**直積**といい，

$$(\Omega_1, P_1)\times(\Omega_2, P_2)$$

であらわす．

乗法律は $\tilde{P}\pi_i^{-1}=P_i\,(i=1, 2)$ をみたす $\tilde{P}$ の中から $P_1\times P_2$ をぬきだす公理であるが，この公理は**エントロピー最大の原理**（これもやはり公理である）から導くことができる．こうしたからといって，公理がなくなったわけでなく，乗法律よりはエントロピー最大の原理の方が人によっては自然な感じがするだけである．一般に P を

$$\Omega = \{a_1, a_2, \cdots, a_m\}$$

の上の確率測度とするとき，

$$\varepsilon(P) = \sum_{i=1}^{m} P\{a_i\}\log\frac{1}{P\{a_i\}}$$

を P の**エントロピー**という．ただし

$$0\log\frac{1}{0} = \lim_{x\to 0} x\log\frac{1}{x} = 0$$

と約束する．今 P_1, P_2 をそれぞれ

$$\Omega_1 = \{a_1, a_2, \cdots, a_l\}, \qquad \Omega_2 = \{b_1, b_2, \cdots, b_m\}$$

の上の確率測度とし，$\tilde{P}$ を $\Omega_1\times\Omega_2$ の上の確率測度で

$$\tilde{P}\pi_1^{-1} = P_1, \qquad \tilde{P}\pi_2^{-1} = P_2$$

をみたすものとする．

$$p_i = P_1\{a_i\}, \qquad q_j = P_2\{b_j\}, \qquad r_{ij} = \tilde{P}\{(a_i, b_j)\}$$

とおくと，上の条件は

$$\sum_{j=1}^{m} r_{ij} = p_i, \qquad i = 1, 2, \cdots, l,$$
$$\sum_{i=1}^{l} r_{ij} = q_j, \qquad j = 1, 2, \cdots, m$$

となる．$\tilde{P}$ のエントロピーは

$$\varepsilon(\tilde{P}) = \sum_{i,j} r_{ij}\log\frac{1}{r_{ij}}$$

である．したがってエントロピー最大の原理から乗法律を導くには，r_{ij} に関する上の付帯条件（p_i, q_j は与えられているとする）のもとで $\varepsilon(\tilde{P})$ を最大にするものが

$$r_{ij} = p_i q_j, \qquad i = 1, 2, \cdots, l;\ j = 1, 2, \cdots, m$$

で与えられることを Lagrange の未定係数法を用いて証明すればよい．その概要をのべておこう．

$$F(r_{11}, r_{12}, \cdots, r_{lm}) = \sum_{i,j} r_{ij}\log\frac{1}{r_{ij}} - \sum_i \alpha_i\Big(\sum_j r_{ij} - p_i\Big) - \sum_j \beta_j\Big(\sum_i r_{ij} - q_j\Big)$$

とおくと

$$\frac{\partial F}{\partial r_{ij}} = \log\frac{1}{r_{ij}} - 1 - \alpha_i - \beta_j = 0,$$

すなわち

$$r_{ij} = e^{-1-\alpha_i-\beta_j}.$$

これと付帯条件 $p_i = \sum_j r_{ij}$, $q_j = \sum_i r_{ij}$ とから

$$p_i = e^{-1}e^{-\alpha_i}\sum_k e^{-\beta_k}, \qquad q_j = e^{-1}e^{-\beta_j}\sum_l e^{-\alpha_l}$$

が得られる．したがって

$$\frac{r_{ij}}{p_i q_j} = \frac{1}{e^{-1}\sum_k e^{-\alpha_k}\sum_l e^{-\beta_l}}.$$

右辺は i, j に無関係で，これを γ とおくと

$$1 = \sum_{i,j} r_{ij} = \gamma \sum_i p_i \sum_j q_j = \gamma$$

となる．故に $r_{ij} = p_i q_j$ となり，望む結果が得られた．

n 個の試行の直結合についても同様の考察が可能で，直結合の確率空間は成分試行のそれの直積である．明らかに

定理 1.10 $\tilde{P} = P_1 \times P_2 \times \cdots \times P_n$ すなわち

$$\tilde{P}\{(\omega_1, \omega_2, \cdots, \omega_n)\} = P_1\{\omega_1\} P_2\{\omega_2\} \cdots P_n\{\omega_n\}$$

ならば

$$\tilde{P}(A_1 \times A_2 \times \cdots \times A_n) = P_1(A_1) P_2(A_2) \cdots P_n(A_n).$$

d) 試行の樹形結合

結合と混合の二つの操作を組み合せて新しい試行をつくることができる．例をあげて，これを説明しよう．

例 1.1 A, B 2人が交互に硬貨を投げ (A から始める)，10回までに A, B のうちで早く表をだした方が勝ちとする．10回までにどちらも表をださなければ，引き分けとする．この試合経過を一つの試行 T と考えて見よう．表がでることを1，裏がでることを0であらわすと，その見本空間 Ω は次の11個の点からなりたっている．

$a_1 = 1$ (A の勝ち)，

$a_2 = 01$ (B の勝ち)，

$a_3 = 001$ (A の勝ち)，

………

$a_{10} = 0000000001$ (B の勝ち)，

$a_{11} = 0000000000$ (引き分け)．

T の確率法則 P が

$$P\{a_k\} = 2^{-k} \quad (k = 1, 2, \cdots, 10), \qquad P\{a_{11}\} = 2^{-10}$$

で与えられることは読者には明らかであろうが，その理由をきちっと説明しようとすると，直結合と混合の2操作を組み合せる必要がおこる．

T の結果 a_1 がでれば，A の勝ちとして勝負は第1回の投げできまったのであるが，かりに10回まで投げ続けるとする．$a_2, a_3, \cdots, a_9$ がでた場合にも同様に10回まで投げ続ける．このようにして得られる‘10回までの投げ続けという試行’を $\tilde{T}$ であらわす．$\tilde{T}$ は1回目の投げ，2回目の投げ，…，10回目の投げの直結合であるから，その確率空間 $(\tilde{\Omega}, \tilde{P})$ は

$$\tilde{\Omega} = \{0, 1\}^{10} = \{(i_1, i_2, \cdots, i_{10}) \mid i_1, i_2, \cdots, i_{10} = 0, 1\},$$
$$\tilde{P}\{(i_1, i_2, \cdots, i_{10})\} = 2^{-10}$$

である．次に $\tilde{\Omega}$ から Ω の上への写像 $X(\tilde{\omega})$ を

$$X(i_1, i_2, \cdots, i_{10}) = \begin{cases} a_1 & (i_1 = 1 \text{ のとき}), \\ a_2 & (i_1 = 0,\ i_2 = 1 \text{ のとき}), \\ \vdots & \\ a_{10} & (i_1 = i_2 = \cdots = i_9 = 0,\ i_{10} = 1 \text{ のとき}), \\ a_{11} & (i_1 = i_2 = \cdots = i_{10} = 0 \text{ のとき}) \end{cases}$$

で定義する．$X(\tilde{\omega})$ は $(\tilde{\Omega}, \tilde{P})$ の上の Ω 値確率変数である．明らかに始めに考えた試行 T は $\tilde{T}$ を $X(\tilde{\omega})$ を通じて混合したものである．したがって T の確率法則 P は X の確率法則 P^X に等しく，$k = 1, 2, \cdots, 10$ に対しては

$$\begin{aligned} P\{a_k\} &= \tilde{P}^X\{a_k\} = \tilde{P}\{X(\tilde{\omega}) = a_k\} \\ &= \tilde{P}\{(i_1, i_2, \cdots, i_{10}) \mid i_1 = i_2 = \cdots = i_{k-1} = 0, i_k = 1\} \\ &= 2^{10-k} \cdot \frac{1}{2^{10}} = 2^{-k}. \end{aligned}$$

同様に

$$P\{a_{11}\} = 2^{-10}.$$

ついで A の勝つ確率を求めると

$$\sum_{k=1}^{5} P\{a_{2k-1}\} = \sum_{k=1}^{5} 2^{-(2k-1)} = \frac{2}{3}(1 - 2^{-10}),$$

同様に B の勝つ確率は

$$\sum_{k=1}^{5} P\{a_{2k}\} = \sum_{k=1}^{5} 2^{-2k} = \frac{1}{3}(1 - 2^{-10}),$$

引き分けの確率はもちろん 2^{-10} である．

また勝負だけに着目した試行 T' を考えると，これは T をさらに混合したもので，その確率空間 (Ω', P') は

$$\Omega' = \{\text{A勝},\ \text{B勝},\ \text{引分}\},$$

$$P'\{\text{A勝}\} = \frac{2}{3}(1-2^{-10}), \qquad P'\{\text{B勝}\} = \frac{1}{3}(1-2^{-10}), \qquad P'\{\text{引分}\} = 2^{-10}$$

で与えられる．

例 1.2 $1, 2, \cdots, n$ の番号札が1枚ずつ合計 n 枚箱 B_0 の中にはいっているとする．まず1枚ぬきだし，残りの $(n-1)$ 枚からもう1枚ぬきだすという試行 T を考えよう．T の確率空間 (Ω, P) が

$$\Omega = \{(i, j) \mid i, j = 1, 2, \cdots, n\ (i \neq j)\}$$

$$P\{(i, j)\} = \frac{1}{n(n-1)}$$

で与えられることは読者には明らかであろうが，その理由も結合，混合の2操作を組み合せて説明することができる．まず与えられた箱 B_0 からのぬきだしを T_0 であらわす．2回目は，T_0 の結果 i の札がでたときには，$(n-1)$ 枚の札

$$1, 2, \cdots, i-1, i+1, \cdots, n$$

の入った箱からのぬきだしである．この箱を B_i であらわし，それからのぬきだしを T_i とかくことにする．この2回のぬきだしを組み合せた T には，$(n+1)$ 個の試行 $T_0, T_1, \cdots, T_n$ が関係し，これらの試行から T を得るには次の樹形図式による．

$$T_0 \begin{cases} 1\text{の札がでると} \longrightarrow T_1 \\ 2\text{の札がでると} \longrightarrow T_2 \\ \cdots\cdots\cdots\cdots\cdots \\ n\text{の札がでると} \longrightarrow T_n \end{cases} \qquad \text{または簡単に} \qquad T_0 \begin{array}{l} \nearrow T_1 \\ \rightarrow T_2 \\ \quad \vdots \\ \searrow T_n \end{array} \qquad \text{とかく．}$$

このような意味で，T を**樹形結合**という．

$T_0, T_1, \cdots, T_n$ の直結合を $\tilde{T}$ とする．これは T とはちがって，T_0 の結果が何であっても $T_1, T_2, \cdots, T_n$ を続いておこなうのである．つまり $(n+1)$ 個の箱 $B_0, B_1, \cdots, B_n$ から1枚ずつぬきだし，その札をならべて得られるベクトル $(i, j_1, j_2, \cdots, j_n)$ を $\tilde{T}$ の結果と考えるのである．さて始めに問題とした試行 T を $\tilde{T}$ に関連させて，次のようにみなすことができる．“B_0 の結果 i のいかんにかかわらず，とにかく $B_1, B_2, \cdots, B_n$ からそれぞれ $j_1, j_2, \cdots, j_n$ をぬきだし，j_i 以外の札は無視して，(i, j_i) を T の結果とする．” こうすると，T は $\tilde{T}$ の混合である．

T_0 の確率空間 (Ω_0, P_0) は

$$\Omega_0 = \{1, 2, \cdots, n\}, \qquad P_0\{j\} = \frac{1}{n}$$

で与えられ，$T_i\ (i=1, 2, \cdots, n)$ の確率空間 (Ω_i, P_i) は

$$\Omega_i = \{1, 2, \cdots, i-1, i+1, \cdots, n\}, \qquad P_i\{j\} = \frac{1}{n-1}$$

である．すべての箱から1枚ずつぬきだす試行 $\tilde{T}$ は $T_0, T_1, T_2, \cdots, T_n$ の直結合 $T_0 \times T_2 \times \cdots \times T_n$ であり，その確率空間 $(\tilde{\Omega}, \tilde{P})$ は

$$\tilde{\Omega} = \Omega_0 \times \Omega_1 \times \cdots \times \Omega_n, \qquad \tilde{P} = P_0 \times P_1 \times \cdots \times P_n$$

で与えられる．$\pi_i : \tilde{\Omega} \to \Omega_i$ を i 射影とすると，$\pi_i(\tilde{\omega})$ は T_i の結果を $(\tilde{\Omega}, \tilde{P})$ の上の関数として表現したものである．したがって T の結果は $(\tilde{\Omega}, \tilde{P})$ の上では関数

$$X(\tilde{\omega}) = (\pi_0(\tilde{\omega}), \pi_{\pi_0(\tilde{\omega})}(\tilde{\omega}))$$

で表現され，T は $\tilde{T}$ を $X(\tilde{\omega})$ で混合したもの $\tilde{T}_X$ と一致する．

以上の考察を背景として，T の確率空間 (Ω, P) を定めよう．

$$\Omega = \{(i, j) \mid i, j = 1, 2, \cdots, n\ (i \neq j)\}$$

は明らかである．$(i, j) \in \Omega$ に対して

$$\begin{aligned} P\{(i, j)\} &= \tilde{P}^X\{(i, j)\} = \tilde{P}\{X(\tilde{\omega}) = (i, j)\} \\ &= \tilde{P}\{\pi_0(\tilde{\omega}) = i, \pi_{\pi_0(\tilde{\omega})}(\tilde{\omega}) = j\} \\ &= \tilde{P}\{\pi_0(\tilde{\omega}) = i, \pi_i(\tilde{\omega}) = j\} \\ &= \tilde{P}\{\pi_0(\tilde{\omega}) = i, \pi_i(\tilde{\omega}) = j, \pi_k(\tilde{\omega}) \in \Omega_k\ (k \neq i, j)\} \\ &= P_0\{i\} P_i\{j\} \prod_{k \neq i, j} P(\Omega_k) \\ &= P_0\{i\} P_i\{j\} \\ &= \frac{1}{n} \cdot \frac{1}{n-1} = \frac{1}{n(n-1)}. \end{aligned}$$

これで (Ω, P) は前に予想した通りになった．——

一般の試行についても同様の考察で**樹形結合**を定義することができる．T_0 の確率空間を

$$(\Omega_0 \equiv \{a_1, a_2, \cdots, a_n\}, P_0)$$

とし，おのおのの $a \in \Omega_0$ に対して試行 T_a が定義され，その確率空間を (Ω_a, P_a) とする．さて樹形結合

$$T_0\begin{cases} a_1 \text{ がでたとき} \longrightarrow T_{a_1} \\ a_2 \text{ がでたとき} \longrightarrow T_{a_2} \\ \cdots\cdots\cdots\cdots\cdots \\ a_n \text{ がでたとき} \longrightarrow T_{a_n} \end{cases}$$

略して

$$\begin{tikzcd}[row sep=small]
 & T_{a_1} \\
T_0 \arrow[ur] \arrow[r] \arrow[ddr] & T_{a_2} \\
 & \vdots \\
 & T_{a_n}
\end{tikzcd}$$

を T としよう．これは上の例で説明したように，直結合 $\tilde{T} = T_0 \times T_{a_1} \times \cdots \times T_{a_n}$ の混合として得られる．$T_a\ (a \in \Omega_0)$ の確率空間を (Ω_a, P_a) とすると，T の見本空間 Ω は当然

$$\Omega = \{(a, b) \mid a \in \Omega_0,\ b \in \Omega_a\}$$

で与えられるが，T の確率法則は次の定理で定まる．

定理 1.11（樹形結合の乗法律）

$$P\{(a, b)\} = P_0\{a\} P_a\{b\}, \qquad (a, b) \in \Omega.$$

証明　$\tilde{T} = T_0 \times T_{a_1} \times T_{a_2} \times \cdots \times T_{a_n}$ の確率空間 $(\tilde{\Omega}, \tilde{P})$ は

$$\tilde{\Omega} = \Omega_0 \times \Omega_{a_1} \times \Omega_{a_2} \times \cdots \times \Omega_{a_n}, \qquad \tilde{P} = P_0 \times P_{a_1} \times P_{a_2} \times \cdots \times P_{a_n}$$

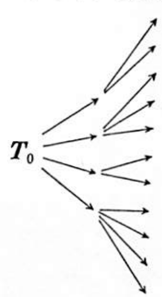
であり，T は $\tilde{T}$ を

$$X(\tilde{\omega}) = (\pi_0(\tilde{\omega}), \pi_{\pi_0(\tilde{\omega})}(\tilde{\omega}))$$

で混合して得られるから，'箱からのぬきだし' と同様にして定理の等式を得る．∎

直結合のときにのべた乗法律は公理であったが，この樹形結合の乗法律は定理として証明されたものである．

さらに複雑な樹形結合も定義できる（下の図式参照）．



このときの乗法律は次の形になる．

$$P\{(a, b, c)\} = P_0\{(a, b, c)\} = P_0\{a\} P_a\{b\} P_{ab}\{c\}.$$

$T_1 = T_2 = \cdots = T_n$ ならば，樹形結合

$$\begin{tikzcd}[row sep=small]
 & T_1 \\
T_0 \arrow[ur] \arrow[r] \arrow[ddr] & T_2 \\
 & \vdots \\
 & T_n
\end{tikzcd}$$

は直結合 $T_0 \times T_1$ と一致する．逆にいえば，直結合 $T_0 \times T_1$ は樹形結合

$$\begin{tikzcd}[row sep=small]
 & T_1 \\
T_0 \arrow[ur] \arrow[r] \arrow[ddr] & T_1 \\
 & \vdots \\
 & T_1
\end{tikzcd}$$

である．したがって樹形結合だけを考えておけば十分で，直結合は特別の場合として含まれてしまう．

例題 1.3　(i)　例 1.2 の試行で 1 回目にぬきだした札の番号を X，2 回目のそれを Y とするとき，

$$P^X,\ P^Y,\ EX,\ EY,\ V(X),\ V(Y),\ \sigma(X),\ \sigma(Y),\ P^{(X,Y)},$$
$$V(X, Y),\ R(X, Y)$$

をもとめよ．

(ii)　箱の中に 1, 2, 3, 4, 5, 6, 7, 8, 9 の番号札が 2 枚ずつ合計 18 枚ある．この中から 4 枚の札をぬきだし，でた順に左から右へならべて 4 けたの数をつくるとき，その数が 5283 より大きくなる確率を求めよ．

［ヒント］　問題の箱から i の札をぬきだす確率を $P_0\{i\}$，この残りから j の札をぬきだす確率を $P_i\{j\}$，さらに残りから k の札をぬきだす確率を $P_{ij}\{k\}$，さらにその残りから l をぬきだす確率を $P_{ijk}\{l\}$ とすると，問題の確率は

$$P_0\{6, 7, 8, 9\} + P_0\{5\} P_5\{3, 4, 6, 7, 9\} + P_0\{5\} P_5\{2\} P_{52}\{9\}$$
$$+ P_0\{5\} P_5\{2\} P_{52}\{8\} P_{528}\{4, 5, 7, 9\}$$

であることに注意せよ．ただし

$$P_i\{A\} = \sum_{j \in A} P_i\{j\}; \qquad P_{ij}, P_{ijk} \text{ についても同様．}$$

(iii)　$\Omega = \{a_1, a_2, \cdots, a_n\}$ の上の確率測度のうちでエントロピー最大のものは $P\{a_i\} \equiv 1/n$ であることを示せ．

§1.4 条件付確率

(Ω, P) を確率空間とし，A, B を Ω の任意の部分集合とし，A は空でないとする．A のもとにおける B の **(条件付) 確率**を

$$P(A\cap B)/P(A)$$

で定義し，$P(B|A)$ または $P_A(B)$ であらわす．$P(A)=0$ のときには $P_A(B)$ は意味がないが，便宜上

$$P_A(B)=\begin{cases}1, & \omega_0\in B,\\ 0, & \omega_0\notin B\end{cases}$$

としておく．ここで ω_0 は A の中の任意に固定した点である．

定理 1.12 (i) $P(A\cap B)=P(A)P_A(B)$, (ii) $P_A(B)$ は，A を固定して B の関数と見たとき，Ω の上の確率測度で，$P_A(A)=1$ である．(この意味で集合関数 P_A を A のもとにおける **(条件付) 確率法則**という.)

証明 $P(A)>0$ のときには定義から直ちにでる．$P(A)=0$ のときにも，上の約束からこの定理がなりたつことがわかる．▌

上の定理の (i) をくりかえして

$$P(A_1\cap A_2\cap A_3)=P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3),$$

一般に

$$P(A_1\cap A_2\cap\cdots\cap A_n)=P(A_1)P_{A_1}(A_2)\cdots P_{A_1\cap A_2\cap\cdots\cap A_{n-1}}(A_n)$$

が得られる．

$P_A(B)$ において，$A=X^{-1}\{x\}$ $(x\in\Omega^X)$ とおいて $P_{X^{-1}\{x\}}(B)$ を定義する．$X^{-1}\{x\}=\{\omega\,|\,X(\omega)=x\}$ であるから，$P_{X^{-1}\{x\}}(B)$ を

$$P_{X=x}(B) \quad \text{または} \quad P(B|X=x)$$

とかく．これは "$X=x$ という条件のもとにおける B の条件付確率" とよばれる．$P_{X=x}(B)$ は $x\in\Omega^X$ と $B\subset\Omega$ との関数である．x を固定したとき，$P_{X=x}(B)$ は B の関数として Ω の上の確率測度となっている．この意味で $P_{X=x}$ を "$X=x$ という条件のもとにおける確率法則" という．次に B を固定して $P_{X=x}(B)$ を x の関数と見ると Ω^X の上の関数である．この x を $X(\omega)$ でおきかえたもの，すなわち

$$P_{X=x}(B)|_{x=X(\omega)}$$

を $P_X(B)$ または $P(B|X)$ であらわし，"$X(\omega)$ の値がわかったときの B の条件付確率" という．B を固定したとき，$P_X(B)$ は $X(\omega)$ の関数，したがって ω の関数であって，(Ω, P) の上の確率変数と考えられる．

定理 1.13 $P(X^{-1}(F)\cap B)=\sum_{x\in F}P\{X=x\}P_{X=x}(B)$

$=E(P_X(B), X^{-1}(F))$, ただし $F\subset\Omega^X$.

証明 $X^{-1}(F)=\sum_{x\in F}X^{-1}\{x\}$

と $P\{X=x\}=P^X\{x\}$ から第1の等式がでる．第2の等式をだすには $P_{X=x}(B)$ を $\varphi_B(x)$ とおいて

$$E^X(P_{X=x}(B), F)=E^X(\varphi_B, F)=E^X(\varphi_B 1_F),$$

$$E(P_X(B), X^{-1}(F))=E(\varphi_B(X(\omega)), X^{-1}(F))=E(\varphi_B(X(\omega))1_F(X(\omega)))$$

を得るが，これが等しいことは定理 1.9 からすぐわかる．▌

P_A は Ω の上の確率法則であるから，Ω の上の関数 $Y(\omega)$ に対して，P_A に関する $Y(\omega)$ の平均値 E_AY を

$$E_A(Y)=\sum_{\omega\in\Omega}Y(\omega)P_A\{\omega\}$$

で定義することができる．これはまた直接に

$$E_A(Y)=\begin{cases}E(Y,A)/P(A), & P(A)>0,\\ Y(\omega_0), & P(A)=0\end{cases}$$

と定義しても同じことである．ここで ω_0 は $P_A(B)$ を $P(A)=0$ のとき定義した際に固定した ω_0 である．したがって常に

$$E(Y,A)=P(A)E_A(Y)$$

となる．$P_{X=x}(B)$, $P_X(B)$ を $P_A(B)$ から定義したと同様にして $E_A(Y)$ から $E_{X=x}(Y), E_X(Y)$ を定義することができる．定理 1.13 に対応して，次の定理を得る．

定理 1.14 $E(Y, X^{-1}(F))=\sum_{X\in F}P\{X=x\}E_{X=x}(Y)=E^X(E_{X=x}(Y), F)$

$=E(E_X(Y), X^{-1}(F))$.

証明 定理 1.13 と同様．▌

$E_A(Y), E_{X=x}(Y), E_X(Y)$ をそれぞれ

$$E(Y|A), \quad E(Y|X=x), \quad E(Y|X)$$

とかくこともある．

$P_A(B)$ で $B=\{\omega\,|\,Y\in F\}=Y^{-1}(F)$ とおいて $P_A\{Y\in F\}$ を定義する．これは

F の関数と見て確率測度である．この確率測度は，$Y(\omega)$ を (Ω, P_A) の上の確率変数と見たときの $Y(\omega)$ の確率法則であるから，$(P_A)^Y$ とかくことができる．同様にして $P_{X=x}\{Y\in F\}, P_X\{Y\in F\}, (P_{X=x})^Y, (P_X)^Y$ が考えられる．

前節にのべた樹形結合は条件付確率と密接な関係がある．樹形結合

$$T_0 \begin{array}{l} \nearrow T_{a_1} \\ \rightarrow T_{a_2} \\ \quad\vdots \\ \searrow T_{a_n} \end{array}$$

を考えよう．T_0 の確率空間を

$$(\Omega_0 \equiv \{a_1, a_2, \cdots, a_n\}, P_0)$$

とし，T_a のそれを (Ω_a, P_a) とする．この樹形結合の確率空間 (Ω, P) は

$$\Omega = \{(a, b) \mid a\in\Omega_0, b\in\Omega_a\},$$

$$P\{(a, b)\} = P_0\{a\}P_a\{b\}$$

で与えられる．T の始めの試行 T_0 の結果，次の試行 ($T_{a_1}, T_{a_2}, \cdots, T_{a_n}$ のうちの一つ) の結果は (Ω, P) の上の確率変数であるが，これを $X(\omega), Y(\omega)$ とする，すなわち

$$X(\omega) = \pi_1(\omega), \qquad Y(\omega) = \pi_2(\omega) \qquad (\pi_i \text{ は射影}).$$

したがって $Y(\omega)$ は $\bigcup_{i=1}^{n}\Omega_{a_i}$ の中の値をとる確率変数であって，

$$P\{(a, b)\} = P\{X=a, Y=b\} = P\{X=a\}P_{X=a}\{Y=b\},$$

$$P\{X=a\} = \sum_{b\in\Omega_a}P\{X=a, Y=b\} = \sum_{b\in\Omega_a}P\{(a, b)\}$$

$$= \sum_{b\in\Omega_a}P_0\{a\}P_a\{b\} = P_0\{a\}.$$

これから

$$P_{X=a}\{Y=b\} = \frac{P\{(a, b)\}}{P\{X=a\}} = \frac{P_0\{a\}P_a\{b\}}{P_0\{a\}} = P_a\{b\}$$

がでる．かくして次の定理を得る．

定理 1.15　$P_{X=a}\{Y=b\} = P_a\{b\}$.　——

同様にして多数回の樹形結合に対しても，その第1回目，第2回目，… の結果を $X, Y, Z, \cdots$ とすれば

$$P_{X=a}\{Y=b\} = P_a\{b\},$$

$$P_{X=a, Y=b}\{Z=c\} = P_{a,b}\{c\},$$

$$\cdots\cdots\cdots\cdots\cdots$$

となる．直結合のときには $P_{X=a}\{Y=b\}$ は a に無関係，$P_{X=a, Y=b}\{Z=c\}$ は a, b に無関係となる．

例題 1.4　$\Omega = A_1 + A_2 + \cdots + A_n$ のとき，次の等式を証明せよ．ただし $P(B) \neq 0$.

$$P_B(A_i) = \frac{P(A_i)P_{A_i}(B)}{\sum_{k=1}^{n}P(A_k)P_{A_k}(B)} \qquad \textbf{(Bayes の定理).}$$

§1.5　独　立　性

これからは確率変数を考えるときには，特に断らない限り，ある定まった確率空間 (Ω, P) の上で考えるものとする．確率変数 X, Y が (互いに) **独立**であるとは

$$P\{X=x, Y=y\} = P\{X=x\}P\{Y=y\}, \qquad x\in\Omega^X,\ y\in\Omega^Y$$

がなりたつことである．このときには

$$P_{X=x}\{Y=y\} = P\{Y=y\} \quad (P(X=x)>0 \text{ のとき}), \qquad y\in\Omega^Y,$$

$$P_{Y=y}\{X=x\} = P\{X=x\} \quad (P(Y=y)>0 \text{ のとき}), \qquad x\in\Omega^X.$$

またこのどの一つの条件からも X, Y の独立性がでる．これらはそれぞれ

$$P_{X=x}\{Y=y\} \text{ が } x \text{ に無関係} \qquad (P(X=x)>0 \text{ のとき}),$$

$$P_{Y=y}\{X=x\} \text{ が } y \text{ に無関係} \qquad (P(Y=y)>0 \text{ のとき})$$

といってもよい．独立という名称はこの事実にもとづいている．

定理 1.16　次の条件は互いに同等である．

(i)　X, Y が独立,

(ii)　$P_{X=x}(Y=y) \neq P(Y=y)$ となる x の値の全体を N とすると，

$$P^X(N) = 0,$$

(iii)　$P_X(Y=y) = P(Y=y)$ a.s.

証明　X, Y が独立とすれば，上に注意したことからすべての $x\in N$ に対して

$$P\{X=x\} = 0.$$

故に

$$P^X(N) = P\{X\in N\} = \sum_{x\in N}P\{X=x\} = 0.$$

これで (i)$\Longrightarrow$(ii) が証明された．さて

$$M = \{\omega \mid P_{X(\omega)}(Y=y) \neq P\{Y=y\}\}$$

とおくと，$\omega \in M$ ならば $X(\omega) \in N$. 故に $M \subset X^{-1}(N)$ (実は $=$ がなりたつ). したがって (ii) から

$$P(M) \leqq P(X^{-1}(N)) = P^X(N) = 0 \quad すなわち \quad P(M) = 0$$

がでるが，これは (iii) を意味する．これで (ii)$\Longrightarrow$(iii) がでた．もし (iii) を仮定すれば，上の M を用いて

$$E(P_{X(\omega)}(Y=y), X^{-1}\{x\}) = E(P_{X(\omega)}(Y=y), X^{-1}\{x\} \setminus M),$$
$$E(P(Y=y), X^{-1}\{x\}) = E(P(Y=y), X^{-1}\{x\} \setminus M).$$

$X^{-1}(x) \setminus M$ の上では $P_{X(\omega)}(Y=y)$ と $P(Y=y)$ とは等しいから，

$$E(P_{X(\omega)}(Y=y), X^{-1}\{x\}) = E(P(Y=y), X^{-1}\{x\}).$$

すなわち定理 1.13 と定理 1.3 の (iii) により

$$P(X=x, Y=y) = P(Y=y)P(X^{-1}\{x\}) = P(X=x)P(Y=y)$$

となり，X, Y の独立性がでる．これで (iii)$\Longrightarrow$(i) が示され，結局 (i), (ii), (iii) の同等性が証明された．▮

定理 1.17 X, Y が独立ならば

$$P(X \in E, Y \in F) = P(X \in E)P(Y \in F).$$

証明 $\{\omega \mid X(\omega) \in E, Y(\omega) \in F\} = \sum_{x \in E} \sum_{y \in F} \{\omega \mid X(\omega) = x, Y(\omega) = y\}$

であるから

$$P\{X \in E, Y \in F\} = \sum_{x \in E} \sum_{y \in F} P\{X=x\}P\{Y=y\} = P(X \in E)P(Y \in F). \quad ▮$$

$X_1, X_2, \cdots, X_n$ が独立であるとは

$$P\{X_1=x_1, X_2=x_2, \cdots, X_n=x_n\} = P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_n=x_n\},$$
$$x_i \in \Omega^{X_i}, \qquad i = 1, 2, \cdots, n$$

がなりたつことである．このときには $X_1, X_2, \cdots, X_n$ の順序を交換したものも独立である．上の式の両辺をすべての $x_n \in \Omega^{X_n}$ について加えると，

$$P\{X_1=x_1, X_2=x_2, \cdots, X_{n-1}=x_{n-1}\}$$
$$= P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_{n-1}=x_{n-1}\}$$

が得られる．これから，独立な確率変数系の部分系はやはり独立であることがわかる．

定理 1.18 $\{X_1, X_2, \cdots, X_n\}$ を二つの部分系

$$\{X_1, X_2, \cdots, X_r\} \quad と \quad \{X_{r+1}, X_{r+2}, \cdots, X_n\}$$

に分ける．$X_1, X_2, \cdots, X_n$ が独立であるためには，次の二つの条件がなりたつことが必要十分である．

(i) 各部分系が独立，すなわち $X_1, \cdots, X_r$ が独立，$X_{r+1}, \cdots, X_n$ が独立，

(ii) 結合変数 $(X_1, \cdots, X_r)$, $(X_{r+1}, \cdots, X_n)$ が独立．

三つ以上の部分系にわけても同様のことがいえる．

証明 下のことに注意すれば，容易に証明できる．

$$X_1 = x_1, \cdots, X_r = x_r, X_{r+1} = x_{r+1}, \cdots, X_n = x_n$$
$$\Longleftrightarrow (X_1, \cdots, X_r) = (x_1, \cdots, x_r), \ (X_{r+1}, \cdots, X_n) = (x_{r+1}, \cdots, x_n). \quad ▮$$

定理 1.19 $X_1, X_2, \cdots, X_n$ が独立ならば

$$P\{X_1 \in E_1, X_2 \in E_2, \cdots, X_n \in E_n\} = P\{X_1 \in E_1\}P\{X_2 \in E_2\}\cdots P\{X_n \in E_n\}.$$

証明 定理 1.17 はこの定理の特別の場合 ($n=2$) である．前定理により，$(X_1, X_2, \cdots, X_{n-1})$ と X_n とは独立であるから，定理 1.17 により

$$左辺 = P\{(X_1, X_2, \cdots, X_{n-1}) \in E_1 \times E_2 \times \cdots \times E_{n-1}, X_n \in E_n\}$$
$$= P\{(X_1, X_2, \cdots, X_{n-1}) \in E_1 \times E_2 \times \cdots \times E_{n-1}\}P\{X_n \in E_n\}$$
$$= P\{X_1 \in E_1, X_2 \in E_2, \cdots, X_{n-1} \in E_{n-1}\}P\{X_n \in E_n\}.$$

この論法をくりかえして，結局これは定理の式の右辺に等しくなることがわかる．▮

上の定理は

$$P^{(X_1, X_2, \cdots, X_n)}(E_1 \times E_2 \times \cdots \times E_n) = P^{X_1}(E_1)P^{X_2}(E_2)\cdots P^{X_n}(E_n),$$

すなわち

$$P^{(X_1, X_2, \cdots, X_n)} = P^{X_1} \times P^{X_2} \times \cdots \times P^{X_n} \qquad (直積測度)$$

を意味する．

定理 1.20 $X_1, X_2, \cdots, X_n$ が独立，$Y_1 = \varphi_1(X_1)$, $Y_2 = \varphi_2(X_2)$, $\cdots$, $Y_n = \varphi_n(X_n)$ ならば，$Y_1, Y_2, \cdots, Y_n$ も独立である．

証明 $$Y_1 = y_1, \ Y_2 = y_2, \ \cdots, \ Y_n = y_n$$
$$\Longleftrightarrow X_1 \in \varphi_1^{-1}(y_1), X_2 \in \varphi_2^{-1}(y_2), \cdots, X_n \in \varphi_n^{-1}(y_n)$$

に注意して前定理を用いよ．▮

集合 $A_1, A_2, \cdots, A_n$ が**独立**であるとは，その指示関数 $1_{A_1}(\omega), 1_{A_2}(\omega), \cdots, 1_{A_n}(\omega)$ が独立なことである．集合は事象をあらわすから，これは事象の独立を定義して

いると考えてよい．

定理 1.21 次の三つの条件は互いに同等である．

(i) $A_1, A_2, \cdots, A_n$ は独立である，

(ii) $A_i'=A_i$ または A_i^c のとき，常に

$$P(A_1'\cap A_2'\cap\cdots\cap A_n') = P(A_1')P(A_2')\cdots P(A_n')$$

がなりたつ，

(iii) 任意の $k=2,3,\cdots,n$ と任意の $1\leqq i_1<i_2<\cdots<i_k\leqq n$ に対し

$$P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}).$$

証明 $e_i(\omega)$ を A_i の指示関数とすれば，(i) は $a_i=0,1\ (i=1,2,\cdots,n)$ に対し

$$P(e_1^{-1}(a_1)\cap e_2^{-1}(a_2)\cap\cdots\cap e_n^{-1}(a_n))$$
$$= P(e_1^{-1}(a_1))P(e_2^{-1}(a_2))\cdots P(e_n^{-1}(a_n))$$

がなりたつことを意味する．$a_i=0,1$ に応じて $e_i^{-1}(a_i)=A_i, A_i^c$ であるから，上の条件は (ii) をかきかえたものにすぎない．故に (i), (ii) は同等である．(i) がなりたてば，$A_1, A_2, \cdots, A_n$ の任意の部分系 $A_{i_1}, A_{i_2}, \cdots, A_{i_k}$ も独立，したがってすでに証明した (i)$\Leftrightarrow$(ii) により，(iii) の等式が得られる．これで (i)$\Longrightarrow$(iii) が証明された．残るのは (iii)$\Longrightarrow$(ii) の証明である．(iii) の式から

$$E(e_{i_1}e_{i_2}\cdots e_{i_k}) = E(e_{i_1})E(e_{i_2})\cdots E(e_{i_k}).$$

さて $(a_1+b_1t_1)(a_2+b_2t_2)\cdots(a_n+b_nt_n)$ を展開したものを

$$\sum_{k,\{i_\nu\}} c_{i_1i_2\cdots i_k}t_{i_1}t_{i_2}\cdots t_{i_k}$$

とすると

$$E((a_1+b_1e_1)(a_2+b_2e_2)\cdots(a_n+b_ne_n)) = \sum_{k,\{i_\nu\}} c_{i_1i_2\cdots i_k}E(e_{i_1}e_{i_2}\cdots e_{i_k}),$$
$$E(a_1+b_1e_1)E(a_2+b_2e_2)\cdots E(a_n+b_ne_n) = \sum_{k,\{i_\nu\}} c_{i_1i_2\cdots i_k}E(e_{i_1})E(e_{i_2})\cdots E(e_{i_k})$$

となるが，上にのべた等式により，右辺は相等しい．故に

$$E((a_1+b_1e_1)(a_2+b_2e_2)\cdots(a_n+b_ne_n))$$
$$= E(a_1+b_1e_1)E(a_2+b_2e_2)\cdots E(a_n+b_ne_n).$$

$A_i'=A_i$ または A_i^c であるから，その指示関数を e_i' とすると，

$$e_i' = e_i \quad \text{または} \quad 1-e_i.$$

したがって e_i' は a_i+be_i の形をしている．故に

$$E(e_1'e_2'\cdots e_n') = E(e_1')E(e_2')\cdots E(e_n'),$$

すなわち

$$P(A_1'\cap A_2'\cap\cdots\cap A_n') = P(A_1')P(A_2')\cdots P(A_n').$$

かくして (iii)$\Longrightarrow$(ii) が証明された．▌

試行 $\boldsymbol{T}_1, \boldsymbol{T}_2, \cdots, \boldsymbol{T}_n$ の直結合を $\boldsymbol{T}$ とし，その確率空間をそれぞれ

$$(\Omega_1, P_1),\quad (\Omega_2, P_2),\quad \cdots,\quad (\Omega_n, P_n),\quad (\Omega, P)$$

とすると，

$$\Omega = \Omega_1\times\Omega_2\times\cdots\times\Omega_n,\qquad P = P_1\times P_2\times\cdots\times P_n$$

となる．$\boldsymbol{T}_i$ の結果が $\omega_i\ (i=1,2,\cdots,n)$ のときには，$\boldsymbol{T}$ の結果 ω は $(\omega_1, \omega_2, \cdots, \omega_n)$ で与えられるから，ω_i を (Ω, P) の上で考えると，

$$\omega_i = \pi_i(\omega) \qquad (\pi_i \text{ は } i \text{ 射影})$$

となる．したがって (Ω_i, P_i) の上の確率変数 $X_i(\omega_i)$ は，(Ω, P) の上では

$$X_i(\omega_i) = X_i(\pi_i(\omega)) = (X_i\circ\pi_i)(\omega)$$

となる．

定理 1.22 (Ω, P) の上で考えると，$\omega_1, \omega_2, \cdots, \omega_n$ は独立，したがって $X_1(\omega_1), X_2(\omega_2), \cdots, X_n(\omega_n)$ も独立である．

証明 $\omega_i=\pi_i(\omega)$ であるから

$$P\{\pi_1(\omega)=a_1, \pi_2(\omega)=a_2, \cdots, \pi_n(\omega)=a_n\}$$
$$= P\{(a_1, a_2, \cdots, a_n)\} = P_1\{a_1\}P_2\{a_2\}\cdots P_n\{a_n\},$$
$$P\{\pi_1(\omega)=a_1\} = P(\{a_1\}\times\Omega_2\times\Omega_3\times\cdots\times\Omega_n)$$
$$= P_1\{a_1\}P_2(\Omega_2)P_3(\Omega_3)\cdots P_n(\Omega_n)$$
$$= P_1\{a_1\}.$$

同様に

$$P\{\pi_i(\omega)=a_i\} = P_i\{a_i\},$$

したがって

$$P\{\pi_1(\omega)=a_1, \pi_2(\omega)=a_2, \cdots, \pi_n(\omega)=a_n\}$$
$$= P\{\pi_1(\omega)=a_1\}P\{\pi_2(\omega)=a_2\}\cdots P\{\pi_n(\omega)=a_n\}.$$

これは $\omega_1=\pi_1(\omega), \omega_2=\pi_2(\omega), \cdots, \omega_n=\pi_n(\omega)$ の独立性を意味する．したがって定理 1.20 により，$X_1(\omega_1), X_2(\omega_2), \cdots, X_n(\omega_n)$ も独立である．▌

例題 1.5 (i) さいをふったとき偶数の目のでることと3の倍数の目のでることとは独立であることを示せ(定理 1.21 (i)$\Leftrightarrow$(iii) の応用)．

(ii) さいを $m+n$ 回ふるとき，始めの m 回にでた目の和と後の n 回にでる目の積とは独立であることを定理 1.18, 1.20 を用いて証明せよ．

(iii) X_1 と X_2 とが独立，(X_1, X_2) と X_3 とが独立，(X_1, X_2, X_3) と X_4 とが独立ならば，X_1, X_2, X_3, X_4 が独立であることを証明せよ．また逆もなりたつことを証明せよ．

(iv) A が A 自身と独立ならば，$P(A)=0$ または 1 であることを示せ ($A\cap A=A$ に注意せよ)．

(v) X が X 自身と独立ならば，X はほとんど確実にある定数に等しいことを証明せよ．

[ヒント] 任意の $a\in\Omega^X$ に対し，$P\{X=a\}$ が 1 または 0 に等しいことを注意せよ．

§1.6 独立な実確率変数

本節では $X, X_1, X_2, \cdots$ はすべて実確率変数をあらわす．

定理 1.23 (平均値の乗法性) $X_1, X_2, \cdots, X_n$ が独立ならば，

$$E(X_1X_2\cdots X_n) = EX_1EX_2\cdots EX_n.$$

証明 $n=1$ のときには当然なりたつ．$n=2$ のときには

$$E(X_1X_2) = \sum x_1x_2P^{(X_1,X_2)}\{(x_1, x_2)\} \qquad (\text{定理 } 1.3\,(\text{v})),$$

ここで $\sum$ は $x_1\in\Omega^{X_1}$, $x_2\in\Omega^{X_2}$ の上にわたる和である．X_1, X_2 が独立であるから

$$P^{(X_1,X_2)}\{(x_1, x_2)\} = P^{X_1}\{x_1\}P^{X_2}\{x_2\},$$

$$EX_i = \sum x_iP^{X_i}\{x_i\} \qquad (\textstyle\sum \text{ は } x_i\in\Omega^{X_i} \text{ の上にわたる})$$

であるから

$$E(X_1X_2) = EX_1EX_2.$$

$X_1, X_2, \cdots, X_n$ が独立であるから，$(X_1, X_2, \cdots, X_{n-1})$ と X_n とが独立 (定理 1.18)．故に $n=2$ の場合を用いて

$$E(X_1X_2\cdots X_{n-1}X_n) = E(X_1X_2\cdots X_{n-1})EX_n.$$

$X_1, X_2, \cdots, X_{n-1}$ は独立系 $X_1, X_2, \cdots, X_n$ の部分系として独立であるから，上の式により，定理が $n-1$ に対してなりたてば，n に対してもなりたつ．▌

平均値の加法性

$$E(X_1+X_2+\cdots+X_n) = EX_1+EX_2+\cdots+EX_n$$

は任意の実確率変数について成立したが，乗法性は独立な場合にだけ成立する．たとえば $X_1=X_2=X$ のときには

$$E(X_1X_2) = EX^2 \geqq (EX)^2 = EX_1EX_2$$

で，$EX^2-(EX)^2=VX$ であるから，上の等号は一般に成立しない．

定理 1.24 (分散の加法性) $X_1, X_2, \cdots, X_n$ の任意の 2 変数が独立ならば

$$V(X_1+X_2+\cdots+X_n) = V(X_1)+V(X_2)+\cdots+V(X_n).$$

証明 $Y_i=X_i-EX_i$ とおくと，Y_i と Y_j $(i\neq j)$ とは独立である．しかも

$$V(X_i) = E(Y_i^2), \qquad V(X_1+X_2+\cdots+X_n) = E((Y_1+Y_2+\cdots+Y_n)^2).$$

平均値の乗法性により

$$E(Y_iY_j) = EY_iEY_j = 0, \qquad i\neq j.$$

したがって

$$E((Y_1+Y_2+\cdots+Y_n)^2) = \sum_i E(Y_i^2)+2\sum_{i<j}E(Y_iY_j) = \sum_i E(Y_i^2),$$

すなわち

$$V(X_1+X_2+\cdots+X_n) = V(X_1)+V(X_2)+\cdots+V(X_n). \qquad ▌$$

$X_1, X_2, \cdots, X_n$ が独立ならば，上の定理の仮定がなりたつから，分散の加法性が得られる．上の定理は $X_1, X_2, \cdots, X_n$ の独立性よりはやや弱い条件でも同じことがいえることを主張している．しかしこれは無条件ではいえない．実際 $X_1=X_2=\cdots=X_n=X$ とおいて見れば，

$$V(X_1+X_2+\cdots+X_n) = n^2V(X), \qquad V(X_1)+V(X_2)+\cdots+V(X_n) = nV(X)$$

となる．

X が実確率変数ならば，任意の $t\in(0,1)$ に対して t^X も実確率変数である．t の関数

$$g^X(t) = E(t^X), \qquad t\in(0,1)$$

を X の**生成関数**という．定理 1.3 (v) により

$$g^X(t) = \sum_{x\in\Omega^X} t^xP^X\{x\}.$$

この式によって，X の確率法則 P^X は生成関数 $g^X(t)$ で定まる．実際 Ω^X の点を大きさの順にならべて

$$x_1 < x_2 < \cdots < x_n$$

とすると，$P^X\{x_k\}$ は次のように順々に定まる．

$$P^X\{x_1\}=\lim_{t\downarrow 0}t^{-x_1}g^X(t),$$

$$P^X\{x_k\}=\lim_{t\downarrow 0}t^{-x_k}\left(g^X(t)-\sum_{i=1}^{k-1}t^{x_i}P^X\{x_i\}\right).$$

$X_1, X_2, \cdots, X_n$ が独立ならば，$t^{X_1}, t^{X_2}, \cdots, t^{X_n}$ が独立であるから，平均値の乗法性により，次の定理が得られる.

定理 1.25（生成関数の乗法性） $X_1, X_2, \cdots, X_n$ が独立とし，$X=X_1+X_2+\cdots+X_n$ とおくと

$$g^X(t)=g^{X_1}(t)g^{X_2}(t)\cdots g^{X_n}(t).$$

例 1.3 $A_1, A_2, \cdots, A_n$ が独立で，かつ

$$P(A_i)=p \qquad (i \text{ に無関係})$$

とする．このとき $A_1, A_2, \cdots, A_n$ のうちでおこるものの回数を N とすると

$$P(N=k)=\binom{n}{k}p^k(1-p)^{n-k}, \qquad k=0,1,2,\cdots,n$$

となることを証明しよう.

A_i の指示関数を $e_i(\omega)$ とすると，

$$g^{e_i}(t)=t^1p(e_i=1)+t^0p(e_i=0)=tp(A_i)+p(A_i{}^c)=tp+(1-p).$$

N を $\{e_i\}$ であらわすと

$$N=e_1+e_2+\cdots+e_n.$$

$\{e_i\}$ は独立であるから，生成関数の乗法性により，

$$g^N(t)=\prod_{i=1}^n g^{e_i}(t)=(tp+(1-p))^n=\sum_{k=0}^n\binom{n}{k}p^k(1-p)^{n-k}t^k,$$

したがって $P(N=k)=P^N\{k\}=\binom{n}{k}p^k(1-p)^{n-k}$.

例 1.4 さいを3回ふったとき，目の和が 9, 10 となる確率 p_9, p_{10} を求めよ.

あらわれる目を X_1, X_2, X_3 とし，その和を S とすると，

$$g^{X_i}(t)=\frac{1}{6}(t+t^2+\cdots+t^6)=\frac{1}{6}(1-t)^{-1}(1-t^6)t.$$

定理 1.22 により X_1, X_2, X_3 は独立であるから，生成関数の乗法性により

$$g^S(t)=\prod_{i=1}^3 g^{X_i}(t)=6^{-3}(1-t)^{-3}(1-t^6)^3t^3$$

$$=6^{-3}\left(1+3t+\frac{3\cdot4}{2}t^2+\frac{4\cdot5}{2}t^3+\frac{5\cdot6}{2}t^4+\cdots\right)(1-3t^6+3t^{12}-t^{18})t^3,$$

$$p_9=\text{“}g^S(t) \text{ の } t^9 \text{ の係数”}=6^{-3}\left(\frac{7\cdot8}{2}-3\right)=\frac{25}{216},$$

$$p_{10}=\text{“}g^S(t) \text{ の } t^{10} \text{ の係数”}=6^{-3}\left(\frac{8\cdot9}{2}-9\right)=\frac{27}{216}.$$

例題 1.6 平均値の乗法性を利用して，“$X_1, X_2, \cdots, X_n$ が独立で，$EX_i=0$, $EX_i{}^2=1$ $(i=1,2,\cdots,n)$ ならば

$$E((X_1+X_1X_2+X_1X_2X_3+\cdots+X_1X_2X_3\cdots X_n)^2)=n\text{”}$$

を証明せよ.

§1.7 大数の法則

さいを多数回ふると，i $(=1,2,\cdots,6)$ の目のでる回数の全回数に対する比（これを i の目の**頻度**という）は大体 1/6 である．これは**大数の法則**とよばれ，誰でも経験していることである．本節の目的はこの経験的事実を数学的に正しく表現し，それを証明しようというのである.

さいをふるという試行の確率空間 (Ω_0, P_0) は

$$\Omega_0=\{1,2,\cdots,6\}, \qquad P_0\{i\}=\frac{1}{6}, \quad i=1,2,\cdots,6$$

で与えられるから，さいを n 回ふるという試行の確率空間 (Ω, P) は

$$\Omega=\Omega_0{}^n, \qquad P=P_0{}^n$$

で与えられる．第 k 回目にでる目を $X_k=X_k(\omega)$ とすると，

$$X_k(\omega)=\pi_k(\omega) \quad (\pi_k \text{ は } k \text{ 射影}), \qquad k=1,2,\cdots,n$$

である．明らかに $X_1, X_2, \cdots, X_n$ は (Ω, P) の上の独立な確率変数で，

$$P\{X_k(\omega)=i\}=\frac{1}{6}, \qquad k=1,2,\cdots,n.$$

さて

$$1_i:\ \Omega_0\longrightarrow\{0,1\}, \qquad 1_i(i)=1, \qquad 1_i(j)=0 \quad (j\neq i)$$

を用いると，n 回目までにでる i の目の頻度 R_i は

$$R_i=R_i(\omega)=\frac{1}{n}\sum_{k=1}^n 1_i(X_k(\omega))$$

で与えられるから，R_i は (Ω, P) の上の確率変数である．(Ω, P) も $R_i(\omega)$ も n

に関係しているが記号を複雑にしないために，n は省略する．大数の法則は，n が十分大きいときには

$$R_i(\omega) \doteqdot \frac{1}{6}, \quad i=1,2,\cdots,6$$

を主張している．

まず問題になるのはこの近似式の意味である．左辺 $R_i(\omega)$ は ω の関数であるから，これを

$$\max_{\omega}\left|R_i(\omega)-\frac{1}{6}\right|<\varepsilon \qquad (\varepsilon \text{ は極めて小さい正数})$$

という意味にとることも考えられる．回数 n をいくら大きくしても，i の目が n 回つづいてでることもあり，そのときには，$R_i(\omega)=1$ となるから，上の解釈は不適当である．ただ，n が大きいときには $R_i(\omega)=1$ のような極端なことがおこる確率は極めて小さい．そこで上の近似式の解釈として，

$$\text{“}P\left\{\left|R_i-\frac{1}{6}\right|>\varepsilon\right\} \text{ が極めて小さい”}$$

をとることが考えられる．これは正しいので，実際証明することができる．もっとはっきりいえば，次のようになる．

定理 1.26 (Bernoulli の大数の法則)　$\varepsilon>0$ に対して $n_0(\varepsilon)$ を十分大きくとれば，すべての $n>n_0(\varepsilon)$ に対し

$$P\left\{\left|R_i-\frac{1}{6}\right|>\varepsilon\right\}<\varepsilon, \quad i=1,2,\cdots,6.$$

証明　i を固定し，$e_k(\omega)=1_i(X_k(\omega))$ とおくと，$X_1, X_2, \cdots, X_n$ の独立性から，$e_1, e_2, \cdots, e_n$ の独立性がでる．分散の加法性により，

$$V(R_i)=V\left(\frac{e_1+e_2+\cdots+e_n}{n}\right)=\frac{1}{n^2}V(e_1+e_2+\cdots+e_n)$$
$$=\frac{1}{n^2}(V(e_1)+V(e_2)+\cdots+V(e_n)),$$

$$E(e_k)=1P\{X_k=i\}+0P\{X_k\neq i\}=\frac{1}{6},$$

$$E(R_i)=E\left(\frac{e_1+e_2+\cdots+e_n}{n}\right)=\frac{E(e_1)+E(e_2)+\cdots+E(e_n)}{n}=\frac{1}{6},$$

$$V(e_k)=E\left(\left(e_k-\frac{1}{6}\right)^2\right)=\left(\frac{5}{6}\right)^2P\{X_k=i\}+\left(\frac{1}{6}\right)^2P\{X_k\neq i\}=\frac{5}{36},$$

$$V(R_i)=\frac{5}{36n}, \qquad \sigma(R_i)=\frac{\sqrt{5}}{6\sqrt{n}}$$

であるから，Čebyšev の不等式により

$$P\left\{\left|R_i-\frac{1}{6}\right|>a\frac{\sqrt{5}}{6\sqrt{n}}\right\}\leqq\frac{1}{a^2}.$$

$a=a(\varepsilon)$ を十分大きくとると，右辺は ε より小さくなり，この a に対して $n_0=n_0(\varepsilon)$ を十分大きくとると，$n>n_0$ なる限り

$$a\frac{\sqrt{5}}{6\sqrt{n}}<a\frac{\sqrt{5}}{6\sqrt{n_0}}<\varepsilon$$

となるから，$n>n_0(\varepsilon)$ なる限り

$$P\left\{\left|R_i-\frac{1}{6}\right|>\varepsilon\right\}<\varepsilon. \qquad \blacksquare$$

例題 1.7　一般の試行のくりかえしについて大数の法則を数学的に表現し，これを証明せよ．

第2章　確率測度

§2.1　一般の試行と確率測度

第1章で有限試行に関して確率論の基本的な考え方を説明した．これから有限試行に限定せず，一般の試行について考えるが，根本的な考え方は全く同じである．

一般の試行 $\boldsymbol{T}$ を考え，その見本空間を Ω とし，確率法則を P とする．見本空間については説明の要はなく，ただ無限集合になるかもしれないだけのことである．確率法則 P も，有限試行の場合と同様に，Ω の部分集合の関数であるが，いくつかの違った点がある．

有限試行の場合には，すべての $A\subset\Omega$ に対して A の中の見本点が試行の結果としてあらわれる確率 $P(A)$ が常に定義されたが，一般にこのようなことは必要でなく，ある種の集合に対してだけ $P(A)$ を定義しておけば十分である．$P(A)$ の定義されるような集合 $A\subset\Omega$ の全体を，集合関数 P の定義域という意味で，$\mathcal{D}(P)$ であらわすことにする．$\mathcal{D}(P)$ は任意の集合族でよいのでなく，次の条件を満足しているものと仮定する．

$$(\sigma.1)\qquad \Omega\in\mathcal{D}(P),$$

$$(\sigma.2)\qquad A\in\mathcal{D}(P)\implies A^c\in\mathcal{D}(P),$$

$$(\sigma.3)\qquad A_n\in\mathcal{D}(P)\ (n=1,2,\cdots)\implies\bigcup_{n=1}^{\infty}A_n\in\mathcal{D}(P).$$

第1の条件は当然であろう．第2の条件は事象 A の確率が定義されておれば，その余事象 A^c の確率も定義されることの要請で，これもまた当然である．第3は $A_1, A_2, \cdots$ のそれぞれの確率が定義されておれば，$A_1, A_2, \cdots$ のどれか少なくとも一つの事象がおこるという事象 $\bigcup_n A_n$ の確率も定義されていることを要請している．これもまた自然な仮定であろう．一般に $(\sigma.1)$, $(\sigma.2)$, $(\sigma.3)$ の三つの条件をみたす集合族を Ω の上の **σ 加法族**とよび，測度論で最も基本的な概念である．

P の値に関しては次の仮定をおく．$P(A)$ とかいたときには，A は $\mathcal{D}(P)$ に属

するものと仮定する.

(P. 1) $P(A) \geqq 0,$

(P. 2) (σ 加法性) $P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n),$

(P. 3) $P(\Omega) = 1.$

第1, 第3の二つの条件は有限試行の場合と同じである. 第2の条件は事象 $A_1, A_2, \cdots$ が互いに排反しているときには, そのどれか少なくとも一つの事象がおこる確率がそれぞれの事象の確率の和に等しいことを要請している. 有限試行の場合には二つの排反事象(したがって有限個の排反事象)に対して同様のことを要請したが, ここでは可算無限個の排反事象に対して要請する. 実際有限試行の場合には Ω が有限集合であるから, $A_1, A_2, \cdots$ が排反しておれば, 当然この中の有限個のもの以外は空集合であるから, σ 加法性を要請して見ても, 加法性と本質的には変らない. (P. 1), (P. 2) をみたす集合関数を**測度**といい, さらに (P. 3) もみたすときには**確率測度**という. $\mathcal{D}(P)$ に属する集合を $\boldsymbol{P}$ **可測集合**といい, $P(A)$ を A の $\boldsymbol{P}$ **測度**(または**測度**)という. 空間 Ω に確率測度 P をそえたものを**確率空間** (Ω, P) という.

試行からはなれて, 一つの確率空間 (Ω, P) が与えられているとする. 今 Ω から1点をぬきだし, しかも $A \in \mathcal{D}(P)$ の中の点がぬきだされる確率が $P(A)$ に等しいようなぬきだしを想定すると, これは一つの試行と見なされる. これを (Ω, P) **からのぬきだし**という. この試行の見本空間はもちろん Ω で, 確率法則は P である. 試行 $\boldsymbol{T}$ の見本空間が Ω で, 確率法則が P であるときには, $\boldsymbol{T}$ は本質的には (Ω, P) からのぬきだしと変らない.

$\Omega=[0, 1]$, $P=$"$[0, 1]$ の上の Lebesgue 測度" とする. P は明らかに Ω の上の確率測度で, $\mathcal{D}(P)$ は Lebesgue 可測集合の全体である. このとき (Ω, P) からのぬきだしをすることを, $[0, 1]$ から**でたらめ**に1点をぬきだすという. ここで P として Lebesgue 測度をとったのは, どの点も一様にぬきだされることからきたので, 点に重みをつけてぬきだすときには,

$$P(E) = \int_E f(\omega)\,d\omega$$

としたらよい. ここに $f(\omega)$ は

$$f(\omega) \geqq 0, \qquad \int_{[0,1]} f(\omega)\,d\omega = 1$$

をみたす Lebesgue 可測関数で, **確率密度**とよばれることがある. このような確率測度は $[0, 1]$ の上の密度 f の確率測度とよばれる. 上の Lebesgue 測度は $f(x) \equiv 1$ の場合である.

$\Omega=[0, 1]$ の上の確率測度といっても, 必ずしも密度をもつとは限らない. 例えば, a を Ω の固定点とし,

$$P(E) = \delta_a(E) = \begin{cases} 1, & a \in E, \\ 0, & a \notin E \end{cases}$$

とすれば, これも Ω の上の確率測度である. このときに $\mathcal{D}(P)=2^\Omega$ ($=\Omega$ のすべての部分集合の族) となっている. (Ω, δ_a) からのぬきだしでは, a がぬきだされる確率は1で $\Omega-\{a\}$ の点がぬきだされる確率は0であるから, このぬきだしはもはやでたらめなぬきだしとはいえないが, ぬきだしには違いない.

以上のように $[0, 1]$ からぬきだす方法はいろいろあるが, 普通**でたらめ**なぬきだしといえば最初にのべた Lebesgue 測度に従ってぬきだす場合をいうのである.

上に1点 a だけぬきだすという極端なぬきだしをあげたが, もう少し一般のものの例として, 有理数だけぬきだす試行の例をあげよう. $\Omega=[0, 1]$ の中の有理数に番号をつけて, $r_1, r_2, \cdots$ とする. 今 P を

$$P(E) = \sum_{i:\, r_i \in E} 2^{-i}$$

とすると, これが Ω の上の確率測度となっていることは容易に証明される(読者試みよ). この場合にも $\mathcal{D}(P)=2^\Omega$ である. Ω の中の有理数全体を $\boldsymbol{Q}_{[0,1]}$ とかくと

$$P(\boldsymbol{Q}_{[0,1]}) = 1, \qquad P(\Omega-\boldsymbol{Q}_{[0,1]}) = 0$$

であるから, (Ω, P) からのぬきだしではほとんど確実に有理数がぬきだされる.

無限試行のもう一つの例として"硬貨を無限回なげる"という試行を考えよう. いつものように表がでることを1, 裏がでることを0とかくと, この試行の見本点は0または1を項とする無限列である. 例えば見本点

$$(0, 0, 1, 0, 1, \cdots)$$

は

裏裏表裏表…

とでることを意味する．したがって見本空間 Ω は

$$\Omega=\{\omega=(\omega_1,\omega_2,\cdots)\,|\,\omega_n=0,1\ (n=1,2,\cdots)\}$$

となる．問題は確率法則 P をこの試行にあうように入れることである．

1回目のなげで表がでるか裏がでるかはそれぞれ確率 1/2 である．これらの事象は Ω の上では，それぞれ

$$\Omega_1=\{\omega=(1,\omega_2,\omega_3,\cdots)\,|\,\omega_n=0,1\ (n=2,3,\cdots)\}$$
$$\Omega_0=\{\omega=(0,\omega_2,\omega_3,\cdots)\,|\,\omega_n=0,1\ (n=2,3,\cdots)\}$$

であらわされるから，

$$P(\Omega_0)=P(\Omega_1)=\frac{1}{2}.$$

同様に1回目，2回目に表表，表裏，裏表，裏裏とでる確率はいずれも確率 1/4 である．したがって

$$\Omega_{ij}=\{\omega=(i,j,\omega_3,\omega_4,\cdots)\,|\,\omega_n=0,1\ (n=3,4,\cdots)\},\qquad i,j=0,1$$

とおくと

$$P(\Omega_{00})=P(\Omega_{01})=P(\Omega_{10})=P(\Omega_{11})=\frac{1}{4}.$$

一般に

$$\Omega_{i_1i_2\cdots i_k}=\{\omega=(i_1,i_2,\cdots,i_k,\omega_{k+1},\omega_{k+2},\cdots)\,|\,\omega_n=0,1\ (n\geqq k+1)\},$$
$$i_1,i_2,\cdots,i_k=0,1$$

に対して

$$P(\Omega_{i_1i_2\cdots i_k})=2^{-k}$$

となる．このようにして得られる集合族

$$\{\Omega_{i_1i_2\cdots i_k}\,|\,i_\nu=0,1\ (\nu=1,2,\cdots,k),\ k=1,2,\cdots\}$$

を $\mathcal{I}$ であらわそう．これで $\mathcal{I}$ の元 I に対しては $P(I)$ が定義された．

この P を拡張して確率測度が得られるならば，その確率測度が求める確率法則である．このためには次節にのべる拡張定理が必要であるから，ここでは立ちいらないことにする(例題 2.2 参照)．次の例題は測度論の復習で，本書を通じてしばしば引用される．

例題 2.1　(i)　実数の集合 $\boldsymbol{R}^1$ の Lebesgue 可測部分集合 E に対して次のように定義された $N_{m,v}$, $C_{m,c}$ はいずれも $\boldsymbol{R}^1$ の上の確率測度であることを示せ．ここに $m\in\boldsymbol{R}^1$, $v>0$, $c>0$ はパラメータである．

Gauss 分布　$$N_{m,v}(E)=\int_E\frac{1}{\sqrt{2\pi v}}e^{-(x-m)^2/2v}dx,$$

Cauchy 分布　$$C_{m,c}(E)=\int_E\frac{c}{\pi}\frac{1}{c^2+(x-m)^2}dx.$$

(ii)　$\mathcal{B}$ を集合 S の上の σ 加法族とする．$\mathcal{B}$ は ϕ, S を含み，可算和，可算交，補，差，上極限，下極限の諸演算で閉じていることを証明せよ．ある集合族がある演算で**閉じている**とは，その集合族に属する集合にその演算を施して得られる集合もその集合族に属することである．

[ヒント]　$\mathcal{B}$ が S を含み，可算無限和，補の二つの演算で閉じていることは σ 加法族の定義そのものである．$\phi\in\mathcal{B}$ は $\phi=S^c$ からでる．したがって $A,B\in\mathcal{B}$ から

$$A\cup B=A\cup B\cup\phi\cup\phi\cup\cdots\in\mathcal{B}$$

がでるので，$\mathcal{B}$ は有限和でも閉じている．したがって $\mathcal{B}$ は可算和で閉じている．他の諸演算で $\mathcal{B}$ が閉じていることをいうには次のことを注意せよ．

$$\bigcap_n A_n=\Big(\bigcup_n A_n^c\Big)^c,\qquad A\setminus B=A\cap B^c,$$
$$\limsup_{n\to\infty}A_n=\bigcap_k\bigcup_{n>k}A_n,\qquad \liminf_{n\to\infty}A_n=\bigcup_k\bigcap_{n>k}A_n.$$

(iii)　確率測度 P は次の性質をもつことを証明せよ．

$P(\phi)=0$,　$P(A^c)=1-P(A)$,　$0\leqq P(A)\leqq 1$,

$P(A+B)=P(A)+P(B)$　(**有限加法性**),

$A\subset B\Longrightarrow P(A)\leqq P(B)$　(**単調性**),　$P(B-A)=P(B)-P(A)$,

A_n が単調列 ならば $P\Big(\lim\limits_{n\to\infty}A_n\Big)=\lim\limits_{n\to\infty}P(A_n)$　(**単調極限定理**),

$P\Big(\liminf\limits_{n\to\infty}A_n\Big)\leqq\liminf\limits_{n\to\infty}P(A_n)\leqq\limsup\limits_{n\to\infty}P(A_n)\leqq P\Big(\limsup\limits_{n\to\infty}A_n\Big)$　(**Fatou の補題**),

$A=\lim\limits_{n\to\infty}A_n$ ならば $P(A)=\lim\limits_{n\to\infty}P(A_n)$　(**連続性**),

$P\Big(\bigcup\limits_n A_n\Big)\leqq\sum\limits_n P(A_n)$　(**劣加法性**),

$P\Big(\bigcap\limits_n A_n\Big)\geqq 1-\sum\limits_n P(A_n^c)$,

$P(A_n)=0\ (n=1,2,\cdots)\Longrightarrow P\Big(\bigcup\limits_n A_n\Big)=0$,

$$P(A_n)=1\ (n=1,2,\cdots)\Longrightarrow P\Big(\bigcap_n A_n\Big)=1.$$

［ヒント］ $\Omega=\Omega+\phi+\phi+\cdots$ から P の σ 加法性により,

$$1=1+P(\phi)+P(\phi)+\cdots.$$

故に $P(\phi)=0$. したがって

$$P(A+B)=P(A+B+\phi+\phi+\cdots)=P(A)+P(B)+0+0+\cdots$$
$$=P(A)+P(B).$$

これから始めの3行の性質はでる．単調極限定理をいうには，まず単調増加ならば,

$$A_1\subset A_2\subset\cdots\Longrightarrow\lim_{n\to\infty}A_n=A_1+(A_2-A_1)+(A_3-A_2)+\cdots$$

に注意し，単調減少ならば，補集合をとって単調増加の場合に帰着せよ．これから Fatou の補題，連続性を導け．劣加法性をいうには

$$\bigcup_n A_n=\sum_n B_n,\qquad B_n=A_n\setminus\bigcup_{i=1}^{n-1}A_i\subset A_n$$

に注意せよ．これから残りの性質を導け．

(iv) P が確率測度のとき，次の不等式を証明せよ．

$$P\Big(\bigcup_n A_n\setminus\bigcup_n B_n\Big)\leqq\sum_n P(A_n\setminus B_n),$$

$$P\Big(\bigcap_n A_n\setminus\bigcap_n B_n\Big)\leqq\sum_n P(A_n\setminus B_n).$$

［ヒント］ $B=\bigcup_n B_n$ とすると，$B\supset B_n\,(n=1,2,\cdots)$，故に

$$\bigcup_n A_n\setminus B=\bigcup_n(A_n\setminus B)\subset\bigcup_n(A_n\setminus B_n).$$

これから第1式を導け．第2式は $A\setminus B=A\cap B^c=B^c\setminus A^c$ を用いて第1式に帰着される．

(v) 次の同等関係を証明せよ．

$x\in\limsup A_n\iff x$ は $A_1,A_2,\cdots$ のうち無限個のものに含まれる,

$x\in\liminf A_n\iff x$ は $A_1,A_2,\cdots$ のうち有限個以外のすべてのものに含まれる．

(vi) $\mathcal{A}$ を S の任意の部分集合族とするとき，$\mathcal{A}$ を含むすべての σ 加法族のうちで最小のもの(これを **$\mathcal{A}$ で生成される σ 加法族**といい，$\sigma[\mathcal{A}]$ であらわす)が

存在することを示せ．

［ヒント］ $\mathcal{A}$ を含むすべての σ 加法族の共通部分をとれ．

(vii) **(Borel 集合族)** S を位相空間とするとき，S の開集合族で生成される σ 加法族を S の **Borel 集合族**といい，$\mathcal{B}(S)$ であらわす．$\mathcal{B}(S)$ の元を S の Borel 集合という．$\boldsymbol{R}^1$ の任意の区間は Borel 集合であることを示せ．また $\mathcal{B}(\boldsymbol{R}^1)$ は有限開区間で生成されることを示せ．

(viii) **(積 σ 加法族)** $\mathcal{B}_1,\mathcal{B}_2,\cdots,\mathcal{B}_n$ をそれぞれ $S_1,S_2,\cdots,S_n$ の上の σ 加法族とするとき，$S_1\times S_2\times\cdots\times S_n$ の部分集合族

$$\{B_1\times B_2\times\cdots\times B_n\mid B_1\in\mathcal{B}_1,B_2\in\mathcal{B}_2,\cdots,B_n\in\mathcal{B}_n\}$$

で生成される $S_1\times S_2\times\cdots\times S_n$ の上の σ 加法族を

$$\mathcal{B}_1\times\mathcal{B}_2\times\cdots\times\mathcal{B}_n$$

であらわし，**積 σ 加法族**という．

$$\mathcal{B}(\boldsymbol{R}^{d_1+d_2+\cdots+d_n})=\mathcal{B}(\boldsymbol{R}^{d_1})\times\mathcal{B}(\boldsymbol{R}^{d_2})\times\cdots\times\mathcal{B}(\boldsymbol{R}^{d_n})$$

を証明せよ．

［ヒント］ $$\mathcal{B}(\boldsymbol{R}^{d_1+d_2})=\mathcal{B}(\boldsymbol{R}^{d_1})\times\mathcal{B}(\boldsymbol{R}^{d_2})$$

をいえば，一般の場合は n についての帰納法で導かれる．

$E_1\subset\boldsymbol{R}^{d_1}$ で $E_1\times\boldsymbol{R}^{d_2}\in\mathcal{B}(\boldsymbol{R}^{d_1+d_2})$ となるものの全体は $\boldsymbol{R}^{d_1}$ の開集合族を含む σ 加法族であるから，$\mathcal{B}(\boldsymbol{R}^{d_1})$ を含む．これは

$$E_1\in\mathcal{B}(\boldsymbol{R}^{d_1})\Longrightarrow E_1\times\boldsymbol{R}^{d_2}\in\mathcal{B}(\boldsymbol{R}^{d_1+d_2})$$

を示す．同様に

$$E_2\in\mathcal{B}(\boldsymbol{R}^{d_2})\Longrightarrow\boldsymbol{R}^{d_1}\times E_2\in\mathcal{B}(\boldsymbol{R}^{d_1+d_2}).$$

したがって

$$E_i\in\mathcal{B}(\boldsymbol{R}^{d_i})\quad(i=1,2)$$
$$\Longrightarrow E_1\times E_2=(E_1\times\boldsymbol{R}^{d_2})\cap(\boldsymbol{R}^{d_1}\times E_2)\in\mathcal{B}(\boldsymbol{R}^{d_1+d_2}).$$

これから

$$\mathcal{B}(\boldsymbol{R}^{d_1})\times\mathcal{B}(\boldsymbol{R}^{d_2})\subset\mathcal{B}(\boldsymbol{R}^{d_1+d_2})$$

がでる．逆の包含関係は次のことからでる．

“$\mathcal{B}(\boldsymbol{R}^d)$ は $(a_1,b_1)\times(a_2,b_2)\times\cdots\times(a_d,b_d)$ なる形のすべての集合で生成される．”

(ix) **(可測写像)** f を S_1 から S_2 の中への写像とし，$\mathcal{B}_1,\mathcal{B}_2$ をそれぞれ S_1,

S_2 上の σ 加法族とする.

$$E \in \mathcal{B}_2 \Longrightarrow f^{-1}(E) \in \mathcal{B}_1$$

のとき, f は**可測** $\mathcal{B}_1/\mathcal{B}_2$ といい, $f \in \mathcal{B}_1/\mathcal{B}_2$ とかく. 可測性に関する次の諸性質を証明せよ.

$$f \in \mathcal{B}_1/\mathcal{B}_2,\ g \in \mathcal{B}_2/\mathcal{B}_3 \Longrightarrow g \circ f \in \mathcal{B}_1/\mathcal{B}_3 \qquad (\text{可測性の\textbf{推移律}}).$$

$\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_n$ がそれぞれ $S_1, S_2, \cdots, S_n$ の上の σ 加法族で, $\pi_i : S_1 \times S_2 \times \cdots \times S_n \to S_i$ が i 射影であれば,

$$\pi_i \in \mathcal{B}_1 \times \mathcal{B}_2 \times \cdots \times \mathcal{B}_n/\mathcal{B}_i, \qquad i = 1, 2, \cdots, n.$$

$\mathcal{B}$ を $S_1 \times S_2 \times \cdots \times S_n$ の上の σ 加法族とするとき,

$$\pi_i \in \mathcal{B}/\mathcal{B}_i \ (i=1, 2, \cdots, n) \Longrightarrow \mathcal{B} \supset \mathcal{B}_1 \times \mathcal{B}_2 \times \cdots \times \mathcal{B}_n.$$

したがって $\mathcal{B}_1 \times \mathcal{B}_2 \times \cdots \times \mathcal{B}_n$ は, $\pi_i \in \mathcal{B}/\mathcal{B}_i \ (i=1, 2, \cdots, n)$ をみたす $S_1 \times S_2 \times \cdots \times S_n$ の上の σ 加法族 $\mathcal{B}$ のうち最小のものである.

(x) f を S_1 から S_2 の中への写像とし, $\mathcal{B}_1$ を S_1 の上の σ 加法族, $\mathcal{A}_2$ を S_2 のある部分集合族とする. もし

$$E \in \mathcal{A}_2 \Longrightarrow f^{-1}(E) \in \mathcal{B}_1$$

ならば, $f \in \mathcal{B}_1/\sigma[\mathcal{A}_2]$.

(xi) $\mathcal{B}_1$ を S_1 の上の σ 加法族, S_2 を位相空間とし, f を S_1 から S_2 への写像とする. $f \in \mathcal{B}_1/\mathcal{B}(S_2)$ のとき, 簡単に $f \in \mathcal{B}_1$ とかき, f は**可測** $\mathcal{B}_1$ または $\mathcal{B}_1$ **可測**という. S_2 の任意の開集合 G に対して $f^{-1}(G) \in \mathcal{B}_1$ ならば, $f \in \mathcal{B}_1$ を示せ. また $S_2 = \boldsymbol{R}$ のとき

$$f^{-1}((-\infty, a)) \in \mathcal{B}_1 \ (a \in \boldsymbol{R}^1) \Longrightarrow f \in \mathcal{B}_1$$

を示せ.

(xii) (**Borel 写像**) S_1, S_2 が位相空間で, S_1 から S_2 の中への写像 f が可測 $\mathcal{B}(S_1)/\mathcal{B}(S_2)$ のとき, f を **Borel 可測写像**または **Borel 写像**という. 連続写像は Borel 可測であることを示せ.

(xiii) (**P 可測**) P を S_1 上の確率測度, S_2 を位相空間, f を S_1 から S_2 への写像とする. $f \in \mathcal{D}(P)$ (すなわち $f \in \mathcal{D}(P)/\mathcal{B}(S_2)$) のとき, f を **P 可測**または単に**可測**という. $A \in \mathcal{D}(P)$ のとき A の指示関数 $1_A : S_1 \to \boldsymbol{R}^1$ は P 可測であることを示せ.

(xiv) $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_n$ をそれぞれ $S, S_1, S_2, \cdots, S_n$ の上の σ 加法族とし, f_i: $S \to S_i \ (i=1, 2, \cdots, n)$, f を $f_1, f_2, \cdots, f_n$ の**積写像**とする, すなわち

$$f:\ S \longrightarrow S_1 \times S_2 \times \cdots \times S_n, \qquad x \longmapsto (f_1(x), f_2(x), \cdots, f_n(x)).$$

このとき

$$f_i \in \mathcal{B}/\mathcal{B}_i \ (i=1, 2, \cdots, n) \Longrightarrow f \in \mathcal{B}/\mathcal{B}_1 \times \mathcal{B}_2 \times \cdots \times \mathcal{B}_n$$

を証明せよ.

[ヒント] $\mathcal{B}' = \{E \subset S_1 \times S_2 \times \cdots \times S_n \mid f^{-1}(E) \in \mathcal{B}\}$

が σ 加法族であること, $E_1 \times E_2 \times \cdots \times E_n \ (E_i \in \mathcal{B}_i,\ i=1, 2, \cdots, n)$ が $\mathcal{B}'$ に含まれることを示せ.

(xv) (**Lebesgue 拡大**) $P(N)=0$ となる P 可測集合を **(P) 零集合**という. P 零集合の部分集合がすべて P 可測 (したがって P 零) であるとき, P を**完備確率測度**という. (P が Ω の上の完備確率測度のとき, 確率空間 (Ω, P) を**完備確率空間**という.) 任意の確率測度 P に対し, 完備拡張は必ず存在し, しかもその中最小の拡張 (これを P の **Lebesgue 拡大**という) があることを証明せよ.

[ヒント]

$$\mathcal{D}(Q) = \{A \mid \text{適当な } B_1, B_2 \in \mathcal{D}(P) \text{ があって } B_1 \subset A \subset B_2,\ P(B_2 - B_1) = 0\}$$

$$Q(A) = P(B_1) \qquad (\text{上の } B_1 \text{ に対し})$$

とすると, Q が P の Lebesgue 拡大となる.

§2.2 確率測度の拡張定理

前節で試行を研究するのに確率測度が重要な役割を演ずることがわかった. 硬貨を無限回投げるという試行の例について説明したように, 見本空間のある種の簡単な集合に対しては, その確率が比較的容易に定義できるが, 問題はこれを確率測度にまで拡張することである. 本節の目的はこの拡張問題を考察することである.

$\mathcal{A}$ を集合 Ω のある部分集合族とする. $\mathcal{A}$ が交演算で閉じているとき, すなわち

$$A, B \in \mathcal{A} \Longrightarrow A \cap B \in \mathcal{A}$$

のとき, $\mathcal{A}$ を**乗法族**という. また $\mathcal{A}$ が可算直和, 固有差で閉じ, しかも Ω を含んでいるとき, $\mathcal{A}$ を **Dynkin 族**という. 任意の部分集合族 $\mathcal{A}$ に対し, $\mathcal{A}$ で生成される σ 加法族 $\sigma[\mathcal{A}]$ を定義したのと同様に $\mathcal{A}$ で生成される Dynkin 族 $\delta[\mathcal{A}]$ を定義することができる. σ 加法族は Dynkin 族であるから, 当然

$$\delta[\mathcal{A}] \subset \sigma[\mathcal{A}]$$

である.

$\mathcal{A}$ が Dynkin 族でかつ乗法族であるとする. $\mathcal{A}$ は Ω を含むから,

$$A \in \mathcal{A} \Longrightarrow A^c = \Omega - A \in \mathcal{A},$$

したがって

$$\begin{aligned} & A_n \in \mathcal{A} \quad (n=1, 2, \cdots) \\ & \Longrightarrow \bigcup_n A_n = \sum_n A_n \cap A_1{}^c \cap A_2{}^c \cap \cdots \cap A_{n-1}{}^c \in \mathcal{A}. \end{aligned}$$

故に $\mathcal{A}$ は σ 加法族となる.

定理 2.1 (Dynkin 族定理) $\mathcal{A}$ が乗法族ならば,

$$\delta[\mathcal{A}] = \sigma[\mathcal{A}].$$

証明 $\delta[\mathcal{A}] \supset \sigma[\mathcal{A}]$ をいえば十分である. これには $\delta[\mathcal{A}]$ が σ 加法族であることをいえばよいが, そのためには $\delta[\mathcal{A}]$ が乗法族であることをいえばよい.

$$\mathcal{D}_1 = \{A \mid \text{すべての } B \in \mathcal{A} \text{ に対し } A \cap B \in \delta[\mathcal{A}]\}$$

とおくと, $\delta[\mathcal{A}]$ が Dynkin 族であることから, $\mathcal{D}_1$ も Dynkin 族となる. また $\mathcal{A}$ が乗法族であるから, $\mathcal{D}_1$ は $\mathcal{A}$ を含む. したがって $\delta[\mathcal{A}]$ の定義により $\delta[\mathcal{A}] \subset \mathcal{D}_1$. これは

$$A \in \delta[\mathcal{A}], \ B \in \mathcal{A} \Longrightarrow A \cap B \in \delta[\mathcal{A}]$$

を意味する. したがって集合族

$$\mathcal{D}_2 = \{B \mid \text{すべての } A \in \delta[\mathcal{A}] \text{ に対し } A \cap B \in \delta[\mathcal{A}]\}$$

は $\mathcal{A}$ を含む. $\mathcal{D}_1$ と同様 $\mathcal{D}_2$ も Dynkin 族であるから, $\delta[\mathcal{A}] \subset \mathcal{D}_2$. これは

$$A \in \delta[\mathcal{A}], \ B \in \delta[\mathcal{A}] \Longrightarrow A \cap B \in \delta[\mathcal{A}],$$

すなわち $\delta[\mathcal{A}]$ が乗法族であることがわかった. ∎

Dynkin 族定理の応用として次の定理が得られる.

定理 2.2 (確率測度に関する一致の定理) P_1, P_2 を Ω の上の確率測度とし, $\mathcal{A}$ は $\mathcal{D}(P_1) \cap \mathcal{D}(P_2)$ に含まれる乗法族とする. P_1, P_2 が $\mathcal{A}$ の上で一致すれば, $\sigma[\mathcal{A}]$ の上で一致する. すなわち

$$P_1(A) = P_2(A), \ A \in \mathcal{A} \Longrightarrow P_1(A) = P_2(A), \ A \in \sigma[\mathcal{A}].$$

証明 集合族

$$\mathcal{B} = \{A \in \mathcal{D}(P_1) \cap \mathcal{D}(P_2) \mid P_1(A) = P_2(A)\}$$

が Dynkin 族であることは, 確率測度の性質からすぐにわかる. 仮定により, $\mathcal{B} \supset \mathcal{A}$, したがって $\mathcal{B} \supset \delta[\mathcal{A}]$. $\mathcal{A}$ が乗法族であるから, 前定理により $\delta[\mathcal{A}] = \sigma[\mathcal{A}]$. 故に $\mathcal{B} \supset \sigma[\mathcal{A}]$. これは P_1, P_2 が $\sigma[\mathcal{A}]$ の上で一致することを意味する. ∎

位相空間の上の確率測度で, その定義域が Borel 集合族と一致するものを **Borel 確率測度**といい, Borel 確率測度の Lebesgue 拡大を**正則確率測度**という. P_1, P_2 を位相空間 S の上の確率測度でその定義域 $\mathcal{D}(P_1), \mathcal{D}(P_2)$ がどれも $\mathcal{B}(S)$ を含むとする. S の開集合族 $\mathcal{G} = \mathcal{G}(S)$ は明らかに乗法族であり, $\mathcal{B}(S)$ は $\mathcal{G}$ で生成される σ 加法族である. 一致の定理により, P_1, P_2 が $\mathcal{G}$ の上で一致すれば, $\mathcal{B}(S)$ の上で一致する. したがって S の上の二つの Borel 確率測度が $\mathcal{G}$ 上で一致すれば, 完全に一致する. また P が正則確率測度のときには, P は, その $\mathcal{B}(S)$ への制限 $P|_{\mathcal{B}(S)}$ (これは Borel 確率測度) の Lebesgue 拡大と一致する. このことから S の上の二つの正則確率測度も, $\mathcal{G}$ の上で一致すれば, 完全に一致することがわかる. 換言すれば

"位相空間上の Borel 確率測度も正則確率測度もその開集合族上の行動で完全に決定される."

上の議論は開集合族のかわりに $\mathcal{B}(S)$ を生成する任意の乗法族 (例えば閉集合族) をとってもなりたつ.

集合 Ω のすべての部分集合 A に対して $\mu^*(A)$ が定義され, 次の 4 条件をみたすとき, 集合関数 μ^* を Ω の上の**外測度**という.

(C. 1) $\quad 0 \leq \mu^*(A) \leq \infty,$

(C. 2) $\quad \mu^*(\phi) = 0,$

(C. 3) (単調性) $\quad A \subset B \Longrightarrow \mu^*(A) \leq \mu^*(B),$

(C. 4) (劣加法性) $\quad \mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$

外測度 μ^* に関して, A が可測 (μ^* 可測) であるとは, 任意の $W \subset \Omega$ に対し

$$\mu^*(W) = \mu^*(W \cap A) + \mu^*(W \cap A^c)$$

となることである. (C. 4) において $A_{m+1} = A_{m+2} = \cdots = \phi$ とおくと, (C. 2) により

$$\mu^*(A_1 \cup A_2 \cup \cdots \cup A_m) \leq \mu^*(A_1) + \mu^*(A_2) + \cdots + \mu^*(A_m)$$

が得られるから,

$$\mu^*(W) \leq \mu^*(W \cap A) + \mu^*(W \cap A^c)$$

は無条件になりたつ．したがって上の可測性の条件は不等式

$$\mu^*(W) \geqq \mu^*(W\cap A)+\mu^*(W\cap A^c)$$

でおきかえてもよい．

定理 2.3（Carathéodory の定理）　μ^* を Ω の上の外測度とし，$\mathcal{M}$ を μ^* 可測な集合の全体とするとき，μ^* を $\mathcal{M}$ に制限したもの $\mu^*|_{\mathcal{M}}$ は Ω の上の測度である．特に $\mu^*(\Omega)=1$ のときには $\mu^*|_{\mathcal{M}}$ は確率測度となる．

証明　$\mu=\mu^*|_{\mathcal{M}}$ とおくと，$\mathcal{D}(\mu)=\mathcal{M}$．まず $\mathcal{M}$ が σ 加法族であることをいう．$\mathcal{M}$ が Ω を含むこと，$\mathcal{M}$ が補演算で閉じていることは容易にわかる．$A, B\in \mathcal{M}$ とすると，

$$\begin{aligned}\mu^*(W) &\geqq \mu^*(W\cap A)+\mu^*(W\cap A^c)\\ &\geqq \mu^*(W\cap A)+\mu^*(W\cap A^c\cap B)+\mu^*(W\cap A^c\cap B^c)\\ &\geqq \mu^*(W\cap (A\cup B))+\mu^*(W\cap (A\cup B)^c)\\ &(\because\ A\cup B=A+A^c\cap B).\end{aligned}$$

これは $A\cup B\in\mathcal{M}$ を意味する．したがって $\mathcal{M}$ は有限和で閉じている．$A_1, A_2, \cdots \in \mathcal{M}$ が互いに素であれば，上と同様に

$$\begin{aligned}\mu^*(W) &\geqq \mu^*(W\cap A_1)+\mu^*(W\cap A_1{}^c)\\ &\geqq \mu^*(W\cap A_1)+\mu^*(W\cap A_1{}^c\cap A_2)+\mu^*(W\cap A_1{}^c\cap A_2{}^c)\\ &\geqq \mu^*(W\cap A_1)+\mu^*(W\cap A_2)+\mu^*(W\cap A_1{}^c\cap A_2{}^c)\\ &\geqq \mu^*(W\cap A_1)+\mu^*(W\cap A_2)+\mu^*(W\cap A_1{}^c\cap A_2{}^c\cap A_3)\\ &\quad +\mu^*(W\cap A_1{}^c\cap A_2{}^c\cap A_3{}^c)\\ &= \mu^*(W\cap A_1)+\mu^*(W\cap A_2)+\mu^*(W\cap A_3)\\ &\quad +\mu^*(W\cap A_1{}^c\cap A_2{}^c\cap A_3{}^c).\end{aligned}$$

これをくりかえして

$$\begin{aligned}\mu^*(W) &\geqq \sum_{n=1}^{N}\mu^*(W\cap A_n)+\mu^*(W\cap A_1{}^c\cap A_2{}^c\cap\cdots\cap A_N{}^c)\\ &\geqq \sum_{n=1}^{N}\mu^*(W\cap A_n)+\mu^*\Big(W\cap \bigcap_{n=1}^{\infty}A_n{}^c\Big) \qquad (\mu^* \text{ の単調性}).\end{aligned}$$

$N\to\infty$ として

$$\begin{aligned}\mu^*(W) &\geqq \sum_{n=1}^{\infty}\mu^*(W\cap A_n)+\mu^*\Big(W\cap\Big(\sum_{n=1}^{\infty}A_n\Big)^c\Big)\\ &\geqq \mu^*\Big(W\cap \sum_{n=1}^{\infty}A_n\Big)+\mu^*\Big(W\cap\Big(\sum_{n=1}^{\infty}A_n\Big)^c\Big),\end{aligned}$$

これは $\sum_{n=1}^{\infty}A_n\in\mathcal{M}$ を意味する．したがって $\mathcal{M}$ は可算直和で閉じている．これから $\mathcal{M}$ が可算和について閉じていることを導くには，

$$\bigcup_n A_n = \sum_{n=1}^{\infty}A_n\cap A_1{}^c\cap A_2{}^c\cap\cdots\cap A_{n-1}{}^c = \sum_{n=1}^{\infty}(A_n{}^c\cup A_1\cup A_2\cup\cdots\cup A_{n-1})^c$$

に注意すればよい．これで $\mathcal{M}$ が σ 加法族であることが証明された．

次に $\mu=\mu^*|_{\mathcal{M}}$ が測度の性質をみたすことをいう．問題になるのは σ 加法性だけである．上に得た不等式で

$$W=\sum_{n=1}^{\infty}A_n \qquad (A_1, A_2, \cdots \in\mathcal{M})$$

とおくと，

$$\mu\Big(\sum_{n=1}^{\infty}A_n\Big)\geqq\sum_{n=1}^{\infty}\mu(A_n)\geqq\mu\Big(\sum_{n=1}^{\infty}A_n\Big).$$

故に

$$\mu\Big(\sum_{n=1}^{\infty}A_n\Big)=\sum_{n=1}^{\infty}\mu(A_n).$$

これで μ が測度であることがわかった．$\mu^*(\Omega)=1$ のときには μ が確率測度となることは明らかである．∎

σ 加法族の定義の三つの条件のうち，σ 加法性（可算和で閉じていること）をゆるめて**加法性**（有限和で閉じていること）でおきかえて**加法族**を定義する．集合 Ω の上の加法族 $\mathcal{A}$ で定義された集合関数 p が

(p. 1)　　$p(A)\geqq 0$,

(p. 2)　（**加法性**）　$p(A+B)=p(A)+p(B)$,

(p. 3)　　$p(\Omega)=1$

をみたすとき，**初等確率測度**という．第1章でのべた有限集合 Ω の上の確率測度は 2^{Ω} を定義域とする初等確率測度である．初等確率測度の性質を導くことは容易であるから，読者にまかせる．σ 加法族は加法族であり，確率測度は初等確率測度であることは定義から明らかであろう．

$[0,1]$ の部分区間の有限直和の全体は加法族であり，そのような集合に対し構成区間の長さの和を対応させると初等確率測度が得られる．また硬貨を無限回投

げる試行（前節）で導入した集合族 $\mathcal{I}$ の中の集合の有限直和の全体 $\mathcal{A}$ も加法族であり，

$$A=\sum_{i=1}^{n} I_i, \quad I_i \in \mathcal{I}$$

に対し

$$P(A)=\sum_{i=1}^{n} P(I_i) \qquad (P(I_i)\ \text{は前節で定義した})$$

とおくと，P は $\mathcal{A}$ の上の初等確率測度となる．

初等確率測度を拡張して確率測度を定義するには次の定理による．

定理 2.4（確率測度の拡張定理） $\mathcal{A}$ を集合 Ω の加法族とし，p を $\mathcal{A}$ の上の初等確率測度とする．p が $\sigma[\mathcal{A}]$ の上の確率測度 P にまで拡張できるための必要十分条件は p が $\mathcal{A}$ の上で σ 加法的なこと，すなわち

$$\text{“}A_1, A_2, \cdots \in \mathcal{A}\ \text{が互いに素で,}\ \sum_{n=1}^{\infty} A_n \in \mathcal{A}\text{”} \Longrightarrow p\left(\sum_{n=1}^{\infty} A_n\right)=\sum_{n=1}^{\infty} p(A_n).$$

がなりたつことである．この拡張 P は一意である．

注意 この条件は

$$\text{“}A_1 \supset A_2 \supset \cdots,\ A_n \in \mathcal{A}\ (n=1,2,\cdots),\ \bigcap_{n=1}^{\infty} A_n=\phi\text{”} \Longrightarrow \lim_{n\to\infty} p(A_n)=0$$

と同等である．これはまた

$$\text{“}A_1 \supset A_2 \supset \cdots,\ A_n \in \mathcal{A}\ (n=1,2,\cdots),\ \inf p(A_n)>0\text{”} \Longrightarrow \bigcap_{n=1}^{\infty} A_n \neq \phi$$

とも同等である．この性質を p の共通点性という．

定理 2.4 の証明 拡張の一意性は確率測度に関する一致の定理からでる．上の条件が拡張可能のために必要なことは明らかであるから，残るのは十分性の証明である．

$$\mu^*(B)=\inf\left\{\sum_{n=1}^{\infty} p(A_n)\ \middle|\ \bigcup_{n=1}^{\infty} A_n \supset B,\ A_n \in \mathcal{A}\ (n=1,2,\cdots)\right\}$$

とおくと，μ^* が外測度となること，しかも $\mu^*(\Omega)=1$ となることは容易にわかる．Carathéodory の定理により，μ^* 可測集合の全体 $\mathcal{M}$ は σ 加法族で，$\mu=\mu^*|_{\mathcal{M}}$ は確率測度である．したがって

$$\mathcal{A} \subset \mathcal{M}, \quad \mu^*(A)=p(A), \quad A \in \mathcal{A}$$

がいえれば，$P=\mu|_{\sigma(\mathcal{A})}$ が求める拡張になる．（$\mathcal{A}\subset\mathcal{M}$ から $\mathcal{D}(\mu)=\mathcal{M}\supset\sigma[\mathcal{A}]$ が

でることに注意せよ.)

p は仮定により σ 加法的であるが，さらに劣 σ 加法的であること，すなわち

$$A \in \mathcal{A},\ A_n \in \mathcal{A}\ (n=1,2,\cdots),\ A \subset \bigcup_{n=1}^{\infty} A_n$$

$$\Longrightarrow A=\sum_{n=1}^{\infty} A \cap \left(A_n \setminus \bigcup_{k=1}^{n-1} A_k\right)$$

$$\Longrightarrow p(A)=\sum_{n=1}^{\infty} p\left(A \cap \left(A_n \setminus \bigcup_{k=1}^{n-1} A_k\right)\right)$$

$$\Longrightarrow p(A) \leqq \sum_{n=1}^{\infty} p(A_n)$$

に注意しよう．これから直ちにすべての $A \in \mathcal{A}$ に対し $p(A) \leqq \mu^*(A)$．また

$$A=A \cup \phi \cup \phi \cup \cdots$$

から

$$\mu^*(A) \leqq p(A)+p(\phi)+p(\phi)+\cdots=p(A)$$

がでるから，結局すべての $A \in \mathcal{A}$ に対し $\mu^*(A)=p(A)$ である．

任意の W に対して

$$W \subset \bigcup_{n=1}^{\infty} A_n$$

となる $A_n \in \mathcal{A}\ (n=1,2,\cdots)$ をとると，任意の $A \in \mathcal{A}$ に対し

$$W \cap A \subset \bigcup_{n=1}^{\infty} A_n \cap A \quad \text{故に} \quad \mu^*(W \cap A) \leqq \sum_{n=1}^{\infty} p(A_n \cap A),$$

$$W \cap A^c \subset \bigcup_{n=1}^{\infty} A_n \cap A^c \quad \text{故に} \quad \mu^*(W \cap A^c) \leqq \sum_{n=1}^{\infty} p(A_n \cap A^c),$$

辺々相加えて

$$\mu^*(W \cap A)+\mu^*(W \cap A^c) \leqq \sum_{n=1}^{\infty} p(A_n).$$

右辺の下限をとると，

$$\mu^*(W \cap A)+\mu^*(W \cap A^c) \leqq \mu^*(W).$$

これは $A \in \mathcal{M}$ を意味する．したがって $\mathcal{A}\subset\mathcal{M}$ となる．これで十分性が証明された．▮

上の拡張定理の応用として Lebesgue-Stieltjes 測度の存在を証明する．今 P を $\boldsymbol{R}^1$ の上の確率測度とし，その定義域が $\mathcal{B}(\boldsymbol{R}^1)$ と一致するとする．このよう

な確率測度を $\boldsymbol{R}^1$ の上の **Borel 確率測度**といい，Borel 確率測度の Lebesgue 拡大を**正則確率測度**という．Borel 確率測度，正測確率測度が一般の位相空間の上でも同様に定義できることはすでにのべた．$\mathcal{B}(\boldsymbol{R}^1)$ を $\mathcal{B}^1$ と略記することが多い．同様に $\mathcal{B}(\boldsymbol{R}^n)$ を $\mathcal{B}^n$ とかく．

さて P を $\boldsymbol{R}^1$ の上のBorel 確率測度とするとき，関数

$$F(x) = F_P(x) = P((-\infty, x]), \qquad x \in \boldsymbol{R}^1$$

を P の**分布関数**という．$F(x)$ が次の3条件をみたすことは確率測度の性質からすぐにわかる．

(F. 1)　(単調性)　$x \leqq y \Longrightarrow F(x) \leqq F(y)$,

(F. 2)　(右連続性)　$F(x+) = F(x) \quad (F(x+) = \lim_{y \downarrow x} F(y))$,

(F. 3)　$F(-\infty) = 0, \quad F(\infty) = 1,$

$$(F(-\infty) = \lim_{x \downarrow -\infty} F(x),\ F(\infty) = \lim_{x \uparrow \infty} F(x)).$$

逆にこの条件を満足する $F(x)$ が与えられたとき，$F(x)$ を分布関数とする正則確率測度 P がただ一つ存在する．この P を $F(x)$ に対応する **Lebesgue-Stieltjes 測度**という．P の存在は上の拡張定理から次のようにして証明される．まず左半開区間 $(a, b]$ に対して

$$p((a, b]) = F(b) - F(a), \qquad -\infty \leqq a < b \leqq \infty$$

と定義する．ただし $(a, \infty]$ は (a, ∞) の意味にとる．左半開区間の有限直和であらわされる集合(これをかりに**初等集合**とよんでおこう)の全体を $\mathcal{A}$ とすると，$\mathcal{A}$ は加法族である．

$$A = \sum_{i=1}^n I_i, \qquad I_i \text{ は左半開区間}$$

に対して

$$p(A) = \sum_{i=1}^n p(I_i)$$

と定義する．$\mathcal{A}$ の元 A を上の形に分解する方法はいく通りもあるが，その分解に応じて上の式で求めた $p(A)$ は分解の方法には無関係に確定する．p が共通点性をもつことをいえば，p は $\mathcal{B}(\boldsymbol{R}^1)$ の上の確率測度に拡張されるが，さらにその Lebesgue 拡大を P とすれば，P が求める確率測度である．p が共通点性をもつことを示そう．

$$A_1 \supset A_2 \supset \cdots, \qquad A_n \in \mathcal{A} \quad (n=1, 2, \cdots), \qquad \alpha = \inf p(A_n) > 0$$

とせよ．F が右連続であるから，任意の有限左半開区間 $I=(a, b]$ に対し，$J=(a+\varepsilon, b]$ をとると，

$$\bar{J} \subset I, \qquad 0 < p(I) - p(J) = F(a+\varepsilon) - F(a) \qquad (\bar{J} = J \text{ の閉包}).$$

F が右連続であるから ε を小さくとることによって，$p(I)-p(J)$ はいくらでも小さくできる．I が無限区間のときにも同様のことがいえるが，条件(F. 3)により，J は有限区間にとることができる．$A_n \in \mathcal{A}$ は左半開区間の有限和であるから，その構成区間のおのおのの I に対して，上の J をとり，その和(直和)を B_n とすると

$$\bar{B}_n \subset A_n.$$

しかも $p(A_n) - p(B_n)$ はいくらでも小さいようにできるから，

$$p(A_n) - p(B_n) < 2^{-n-1}\alpha, \qquad n = 1, 2, \cdots$$

となるように $\{B_n\}$ を定める．

$$p(A_n) - p\left(\bigcap_{i=1}^n B_i\right) = p\left(\bigcup_{i=1}^n (A_n - B_i)\right) \leqq p\left(\bigcup_{i=1}^n (A_i - B_i)\right)$$

$$\leqq \sum_{i=1}^n p(A_i - B_i) < \sum_{i=1}^n 2^{-i-1}\alpha < \frac{\alpha}{2},$$

$p(A_n) \geqq \alpha$ であるから，

$$p\left(\bigcap_{i=1}^n B_i\right) > \frac{\alpha}{2},$$

したがってもちろん

$$\bigcap_{i=1}^n B_i \neq \phi.$$

故に

$$\bigcap_{i=1}^n \bar{B}_i \neq \phi, \qquad n = 1, 2, \cdots.$$

$\bar{B}_i$ は有界閉集合であるから，Cantor の共通点定理により

$$\bigcap_{n=1}^\infty \bar{B}_n \neq \phi.$$

$A_n \supset \bar{B}_n$ であるから，$\bigcap_{n=1}^\infty A_n \neq \phi$ がでて，p の共通点性が検証されたことになる．

これで拡張 P の存在はわかったが，一意性は確率測度に関する一致の定理か

ら明らかである．（左半開区間の全体は乗法族であり，しかも Borel 集合族を生成することに注意せよ．）

例題 2.2 前節の硬貨を無限回投げる試行に関して導入した集合関数 $P(I)$, $I\in\mathscr{I}$ は Ω の上の確率測度にまで拡張できることを証明せよ．

［ヒント］ $\mathscr{I}$ の元の有限直和であらわされる集合の全体 $\mathscr{A}$ は加法族であり，P を $\mathscr{A}$ の上まで自然な方法で拡張したものは初等確率測度である．したがってこの初等確率測度 P が共通点性をもつことをいえばよい．今

$$\varphi\colon \Omega \longrightarrow \boldsymbol{K} \quad (\text{Cantor 集合}), \qquad (\omega_1, \omega_2, \cdots) \longmapsto \sum_{n=1}^{\infty}\frac{2\omega_n}{3^n}$$

を考えると，φ は1対1対応である．$A\in\mathscr{A}$ に対して $\varphi(A)$ は $\boldsymbol{K}$ の閉部分集合であり，$\boldsymbol{K}$ はコンパクトであるから，P の共通点性は Cantor の共通点定理からでる．

§2.3 確率測度の直積

第1章で，有限集合の上の確率測度の直積が，直結合試行や独立確率変数に関連してあらわれた．この事情は一般の試行に対しても同様である．

まず二つの確率測度の直積の定義から始めよう．P_1, P_2 をそれぞれ Ω_1, Ω_2 の上の確率測度とし，

$$\Omega = \Omega_1\times\Omega_2$$

とする．積 σ 加法族 $\mathcal{D}(P_1)\times\mathcal{D}(P_2)$ を定義域とする Ω の上の確率測度 P で

$$P(B_1\times B_2) = P_1(B_1)P_2(B_2), \qquad B_i\in\mathcal{D}(P_i), \qquad i=1,2$$

をみたすものを P_1, P_2 の**直積**といい，$P_1\times P_2$ であらわし，確率空間 (Ω, P) を (Ω_1, P_1) と (Ω_2, P_2) との直積といい，

$$(\Omega, P) = (\Omega_1, P_1)\times(\Omega_2, P_2)$$

であらわす．例えば正方形 $[0,1]^2$ の上の Lebesgue 測度は $[0,1]$ の上の Lebesgue 測度とそれ自身との直積である．第1章では $P=P_1\times P_2$ を

$$P\{(\omega_1, \omega_2)\} = P_1\{\omega_1\}P_2\{\omega_2\}$$

で定義したが，この定義では都合が悪いことは，上の Lebesgue 測度の例からも明らかであろう．

まず問題になるのは直積の存在と一意性である．Ω の部分集合で

$$B_1\times B_2, \qquad B_i\in\mathcal{D}(P_i), \qquad i=1,2$$

の形のものの全体 $\mathscr{I}$ は乗法族であり，しかも $\mathcal{B}\equiv\mathcal{D}(P_1)\times\mathcal{D}(P_2)$ を生成するから，確率測度に関する一致の定理により，直積の一意性は明らかである．存在を証明するには，前節の拡張定理を用いる．まず

$$p(B_1\times B_2) = P_1(B_1)P_2(B_2)$$

により $\mathscr{I}$ の上で集合関数 p を定義し，これを

$$\mathscr{A} = \left\{A=\sum_{i=1}^{n} I_i \,\middle|\, I_i\in\mathscr{I}, n=1,2,\cdots\right\} \qquad (\text{これは } \Omega \text{ の上の加法族である})$$

の上の初等確率測度にまで拡張する．p が共通点性を持つことを示せば，p は $\sigma[\mathscr{A}](=\mathcal{D}(P_1)\times\mathcal{D}(P_2))$ の上の確率測度にまで拡張され，この拡張が求める確率測度である．

$\mathscr{A}$ の元 A は

$$A = \sum_{i=1}^{n} B_{1,i}\times B_{2,i},$$

$$B_{1,i}\in\mathcal{D}(P_1), \qquad B_{2,i}\in\mathcal{D}(P_2), \qquad i=1,2,\cdots,n$$

の形にかけることは当然であるが，さらに $B_{1,i}$ $(i=1,2,\cdots,n)$ が互いに素で，しかもその和が Ω_1 となるようにできる．ここで $B_{2,i}$ の中には空のものもあり得る．（図2.1参照．）このとき，もし

$$p(A) \geqq \alpha > 0$$

ならば，

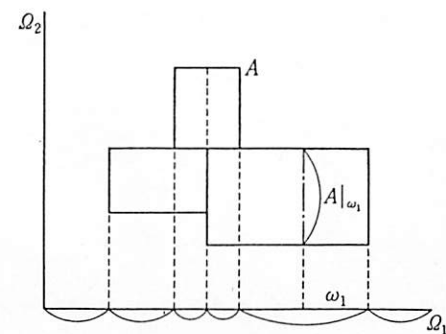


図2.1

$$A_1(\alpha)=\sum B_{1,i},\qquad P_2(B_{2,i})\geqq\frac{\alpha}{2}$$

の P_1 測度は $\alpha/2$ 以上である．上の $\sum$ の意味は $P_2(B_{2,i})\geqq\alpha/2$ をみたすような i についてだけ加えることを意味する．番号をつけかえて，この i の全体を $1,2,\cdots,r$ とすると，

$$\begin{aligned}\alpha\leqq p(A)&=\sum_{i=1}^{r}P_1(B_{1,i})P_2(B_{2,i})+\sum_{i=r+1}^{n}P_1(B_{1,i})P_2(B_{2,i})\\&\leqq\sum_{i=1}^{r}P_1(B_{1,i})\cdot 1+\sum_{i=r+1}^{n}P_1(B_{1,i})\frac{\alpha}{2}\\&\leqq P_1(A_1(\alpha))+P_1(A_1(\alpha)^c)\frac{\alpha}{2}\\&\leqq P_1(A_1(\alpha))+\frac{\alpha}{2}\end{aligned}$$

となり，

$$P_1(A_1(\alpha))\geqq\frac{\alpha}{2}$$

となる．このことを頭において，p が共通点性をもつことを証明しよう．

さて $\mathscr{A}$ の中の減少列

$$A'\supset A''\supset\cdots\supset A^{(n)}\supset\cdots$$

が $p(A^{(n)})\geqq\alpha>0$ をみたすとする．このとき

$$A_1'(\alpha)\supset A_1''(\alpha)\supset A_1'''(\alpha)\supset\cdots$$

であることは明らかである．上の注意から，これらの集合の P_1 測度は $\alpha/2$ 以上である．したがって

$$P_1\left(\bigcap_{n=1}^{\infty}A_1^{(n)}(\alpha)\right)=\lim_{n\to\infty}P_1(A_1^{(n)}(\alpha))\geqq\frac{\alpha}{2}.$$

故にこの交わりは空ではない．その中から1点 ω_1 をとると，

$$A^{(n)}|_{\omega_1}=\{\omega_2\in\Omega_2\,|\,(\omega_1,\omega_2)\in A^{(n)}\}\qquad(A^{(n)}\text{ の }\omega_1\text{ による}\textbf{切口})$$

の P_2 測度は $\alpha/2$ 以上であることは $A_1^{(n)}(\alpha)$ の定義から明らかである．しかも $\{A^{(n)}\}$ の減少性により $\{A^{(n)}|_{\omega_1}\}$ も減少する．したがって上と同様にこの集合列は共通点をもつ．その一つを ω_2 とすれば，(ω_1,ω_2) は $\{A^{(n)}\}$ の共通点である．これで p の共通点性が示された．

三つ以上の確率測度の直積は

$$P_1\times P_2\times P_3=(P_1\times P_2)\times P_3,$$
$$P_1\times P_2\times P_3\times P_4=(P_1\times P_2\times P_3)\times P_4,$$
$$\cdots\cdots\cdots\cdots\cdots$$

で定義できる．明らかに $P_1\times P_2\times\cdots\times P_n$ は

$$\mathscr{D}(P_1\times P_2\times\cdots\times P_n)=\mathscr{D}(P_1)\times\mathscr{D}(P_2)\times\cdots\times\mathscr{D}(P_n),$$
$$(P_1\times P_2\times\cdots\times P_n)(B_1\times B_2\times\cdots\times B_n)=P_1(B_1)P_2(B_2)\cdots P_n(B_n),$$
$$B_i\in\mathscr{D}(P_i),\qquad i=1,2,\cdots,n$$

で定まる．$P_1\times P_2\times\cdots\times P_n$ を $\prod_{i=1}^{n}P_i$ とかくこともある．有限個の確率空間の直積も

$$\prod_{i=1}^{n}(\Omega_i,P_i)=\left(\prod_{i=1}^{n}\Omega_i,\ \prod_{i=1}^{n}P_i\right)\qquad\left(\prod_{i=1}^{n}\Omega_i=\Omega_1\times\Omega_2\times\cdots\times\Omega_n\right)$$

で定義される．

次に可算個の確率測度の直積を考えよう．$P_1,P_2,\cdots$ をそれぞれ $\Omega_1,\Omega_2,\cdots$ の上の確率測度とし，

$$\Omega=\prod_{n=1}^{\infty}\Omega_n=\Omega_1\times\Omega_2\times\cdots\times\Omega_n\times\cdots$$

とする．$\pi_n:\Omega\to\Omega_n$ を n 射影とする．Ω の上の集合族

$$\pi_n^{-1}(B_n),\qquad B_n\in\mathscr{D}(P_n),\qquad n=1,2,\cdots$$

で生成される σ 加法族を

$$\prod_{n=1}^{\infty}\mathscr{D}(P_n)\quad\text{または}\quad\mathscr{D}(P_1)\times\mathscr{D}(P_2)\times\cdots$$

であらわし，$\mathscr{D}(P_n)$ $(n=1,2,\cdots)$ の**積 σ 加法族**という．この積加法族の定義は任意の σ 加法族の列に対しても通用する．さて Ω の上の確率測度 P の定義域が $\prod_{n=1}^{\infty}\mathscr{D}(P_n)$ で，しかも

$$P\left(\bigcap_{i=1}^{n}\pi_i^{-1}(B_i)\right)=\prod_{i=1}^{n}P_i(B_i),$$
$$B_i\in\mathscr{D}(P_i),\qquad i=1,2,\cdots,n,\qquad n=1,2,\cdots$$

がなりたつとき，P を P_n $(n=1,2,\cdots)$ の**直積確率測度**といい，

$$P=\prod_{n=1}^{\infty}P_n\quad\text{または}\quad P_1\times P_2\times\cdots$$

であらわす．Ω の部分集合で

$$\bigcap_{i=1}^{n}\pi_i^{-1}(B_i) \quad すなわち \quad B_1\times B_2\times\cdots\times B_n\times\Omega_{n+1}\times\Omega_{n+2}\times\cdots$$

の形にかけるものの全体を $\mathscr{I}$ とする．直積 P は $\mathscr{I}$ の上では定義によってその値が与えられている．しかも $\mathscr{I}$ は乗法族であり，$\sigma[\mathscr{I}]$ は P の定義域 $\prod_{n=1}^{\infty}\mathscr{D}(P_n)$ と一致するから，確率測度一致の定理により，直積 P は存在するとしてもただ一つである．

さて直積 P の存在を拡張定理を用いて証明しよう．$\mathscr{I}$ の元の有限直和であらわされる集合の全体を $\mathscr{A}$ とすると，$\mathscr{A}$ は加法族であり，$\mathscr{D}(P)\equiv\prod_{n=1}^{\infty}\mathscr{D}(P_n)$ を生成する．さて $\mathscr{I}$ の元に対しては

$$P(B_1\times B_2\times\cdots\times B_n\times\Omega_{n+1}\times\Omega_{n+2}\times\cdots)=P_1(B_1)P_2(B_2)\cdots P_n(B_n)$$

と定義することは当然である．ここでこまかい注意をしておくと，

$$\begin{aligned}&P(B_1\times B_2\times\cdots\times B_n\times\Omega_{n+1}\times\Omega_{n+2}\times\cdots)\\&=P_1(B_1)P_2(B_2)\cdots P_n(B_n)P_{n+1}(\Omega_{n+1})\end{aligned}$$

といってもよいはずであるが，$P_{n+1}(\Omega_{n+1})=1$ であるから，どちらにしても同じである．ここで P_n が確率測度ということがきいている．この $\mathscr{I}$ の上で定義された P は自然な方法で $\mathscr{A}$ の上の初等確率測度まで拡張される．残るのはこれをさらに拡張して $\sigma[\mathscr{A}]$ すなわち $\prod_{n=1}^{\infty}\mathscr{D}(P_n)$ の上の確率測度にまで拡張することであるが，このためには P が共通点性をもつことを証明すればよい．

$A^{(1)}, A^{(2)},\cdots$ を $\mathscr{A}$ の中の減少列とし，$P(A^{(n)})>\alpha>0\ (n=1,2,\cdots)$ とする．このときすべての $A^{(n)}$ に属する点 $\omega=(\omega_1,\omega_2,\cdots)\in\Omega$ が存在することをいえば，P が共通点性をもつ．$\Omega_1\times\Omega_2\times\cdots$ の上の加法族 $\mathscr{A}$ および初等確率測度 P を定めたのと同様に $\Omega'=\Omega_2\times\Omega_3\times\cdots$ の上で $\mathscr{A}'$ と P' を定めると，任意の $A\in\mathscr{A}$ は

$$A=\sum_{i=1}^{n}B_{1,i}\times B_{2,i},\quad B_{1,i}\in\mathscr{D}(P_1),\quad B_{2,i}\in\mathscr{A}',\quad \Omega_1=\sum_{i=1}^{n}B_{1,i}$$

の形にかけて，

$$P(A)=\sum_{i=1}^{n}P_1(B_{1,i})P'(B_{2,i})$$

となることに注意すると，二つの確率測度の直積を定めたときの論法を利用して，$\omega_1\in\Omega_1$ を定め

$$P'(A^{(n)}|_{\omega_1})\geqq\frac{\alpha}{2},\quad n=1,2,\cdots,$$

したがって

$$A^{(n)}|_{\omega_1}\neq\phi,\quad n=1,2,\cdots$$

とすることができる．ここに $A^{(n)}|_{\omega_1}$ は $A^{(n)}$ の ω_1 による**切口**で，

$$A^{(n)}|_{\omega_1}=\{(\omega_2,\omega_3,\cdots)\,|\,(\omega_1,\omega_2,\cdots)\in A^{(n)}\}\in\mathscr{A}'$$

で定められる．$A^{(n)}|_{\omega_1}\ (n=1,2,\cdots)$ は $\Omega_2\times\Omega_3\times\cdots$ の中の減少集合列であり，その P' 測度はすべて $\alpha/2$ 以上である．同じ論法を $A^{(n)}|_{\omega_1}$ に適用して ω_2 を定め

$$P''(A^{(n)}|_{(\omega_1,\omega_2)})\geqq\frac{\alpha}{4},$$

したがって

$$A^{(n)}|_{(\omega_1,\omega_2)}\neq\phi,\quad n=1,2,\cdots$$

とすることができる．ここで P'' は $\Omega_2''=\Omega_3\times\Omega_4\times\cdots$ の上の初等確率測度で，P, P' と同様に定義され，$A^{(n)}|_{(\omega_1,\omega_2)}$ は $A^{(n)}$ の (ω_1,ω_2) による切口で

$$A^{(n)}|_{(\omega_1,\omega_2)}=\{(\omega_3,\omega_4,\cdots)\,|\,(\omega_1,\omega_2,\omega_3,\omega_4,\cdots)\in A^{(n)}\}$$

と定義される．これをくりかえして $\omega_k\ (k=1,2,\cdots)$ を定め，

$$A^{(n)}|_{(\omega_1,\omega_2,\cdots,\omega_k)}\neq\phi,\quad n=1,2,\cdots,\quad k=1,2,\cdots$$

とすることができる．

$\omega=(\omega_1,\omega_2,\cdots)$ とおくと，

$$\omega\in A^{(n)},\quad n=1,2,\cdots$$

であることを証明しよう．$A^{(n)}\in\mathscr{A}$ であるから，$A^{(n)}$ は $\mathscr{I}$ の元の有限直和である．したがって $A^{(n)}$ は

$$A^{(n)}=B\times\Omega_{N+1}\times\Omega_{N+2}\times\cdots,\quad B\subset\Omega_1\times\Omega_2\times\cdots\times\Omega_N$$

の形をしている．N はもちろん n に関係する．$\omega_1,\omega_2,\cdots$ の定め方により

$$A^{(n)}|_{(\omega_1,\omega_2,\cdots,\omega_N)}\neq\phi,$$

すなわち

$$(\omega_1,\omega_2,\cdots,\omega_N)\in B.$$

したがって任意の $\omega_{N+1}',\omega_{N+2}',\cdots$ に対し

$$(\omega_1,\omega_2,\cdots,\omega_N,\omega_{N+1}',\omega_{N+2}',\cdots)\in B\times\Omega_{N+1}\times\Omega_{N+2}\times\cdots=A^{(n)}$$

であり，特に

$$\omega=(\omega_1,\omega_2,\cdots,\omega_N,\omega_{N+1},\omega_{N+2},\cdots)\in A^{(n)}.$$

これで $\mathscr{A}$ の上の初等確率測度 P が共通点性をもつことが示され，したがって直

積 $\prod_{n=1}^{\infty} P_n$ の存在が示されたことになる.

直積測度は一般の測度に対しても考えられる. m を Ω の上の測度とする. $m(\Omega)<\infty$ のとき, **有限測度**といい,

$$\Omega=\sum_{n=1}^{\infty} A_n, \quad m(A_n)<\infty, \quad n=1,2,\cdots$$

なる $\{A_n\}$ が存在するとき, **σ 有限測度**という. m が有限測度ならば, $P(A)=m(A)/m(\Omega)$ ($m(\Omega)>0$ のとき) が確率測度となることを用いて, 有限測度の直積は確率測度の直積に帰着される. すなわち

$$m_1\times m_2=m_1(\Omega_1)m_2(\Omega_2)\frac{m_1}{m_1(\Omega_1)}\times\frac{m_2}{m_2(\Omega_2)},$$

($m_1(\Omega_1)=0$ または $m_2(\Omega_2)=0$ のときには $m_1\times m_2$ は恒等的に 0 であるとする) とおけば

$$(m_1\times m_2)(A_1\times A_2)=m_1(A_1)m_2(A_2)$$

となり, $m_1\times m_2$ が m_1, m_2 の直積の性質をもっている. このような $m_1\times m_2$ の一意性は一致の定理からすぐにでる. この考えで有限個の有限測度の直積も定義できるが, 無限個のときに

$$\prod_{n=1}^{\infty} m_n=\prod_{n=1}^{\infty} m_n(\Omega_n)\prod_{n=1}^{\infty}\frac{m_n}{m_n(\Omega_n)}$$

と定義し得るためには, $\prod_{n=1}^{\infty} m_n(\Omega_n)$ が意味をもつ必要がある. この無限積が絶対収束するには

$$\sum_{n=1}^{\infty}|m_n(\Omega_n)-1|<\infty$$

がなりたつことが必要十分であることは明らかである.

m_1, m_2 がそれぞれ Ω_1, Ω_2 の上の σ 有限測度であれば,

$$\Omega_i=\sum_{n=1}^{\infty} A_{in}, \quad m_i(A_{in})<\infty, \quad n=1,2,\cdots, \ i=1,2$$

なる $\{A_{in}\}$ が存在する. したがって

$$m_{in}(B_i)=m_i(B_i\cap A_{in}), \quad i=1,2$$

は Ω_i の上の有限測度となり, $m_{1n}\times m_{2k}$ が定まる. $m_1\times m_2$ を

$$(m_1\times m_2)(A)=\sum_{n,k=1}^{\infty}(m_{1n}\times m_{2k})(A)$$

と定義すれば, $m_1\times m_2$ は直積測度の性質

$$(m_1\times m_2)(A_1\times A_2)=m_1(A_1)m_2(A_2)$$

をもっている. 同様に $m_1\times m_2\times\cdots\times m_n$ も定義される.

試行列 $\boldsymbol{T}_n$ ($n=1,2,\cdots$)(有限または可算無限)が与えられたとき, その直結合 $\boldsymbol{T}=\boldsymbol{T}_1\times\boldsymbol{T}_2\times\cdots$ の見本空間 Ω は成分試行の見本空間 $\Omega_1, \Omega_2, \cdots$ の直積であることは, 有限試行の場合と同じであるが, $\boldsymbol{T}$ の確率法則 P は

$$P=P_1\times P_2\times\cdots, \quad P_n \text{ は } \boldsymbol{T}_n \text{ の確率法則}$$

で与えられる. これは有限試行の場合と同様, 公理として認めるべきものである. したがって直結合を考えることは成分試行の確率空間の直積を考えることに帰着される.

例題 2.3 (i) さいを無限につづけてふったとき, n 回目に始めて 6 の目がでる確率 p_n を求めよ.

[ヒント] まずこの試行の確率空間 (Ω, P) を定め, 次に問題の事象をあらわす Ω の部分集合 A_n を定め, $p_n=P(A_n)$ を計算せよ. このような簡単な問題でも, 厳密に解くためには, 確率測度の無限直積の概念が必要になることに注意すべきである.

(ii) 確率測度 P の Lebesgue 拡大を $\bar{P}$ とかくとき,

$$\overline{P_1\times P_2\times\cdots}=\overline{\bar{P}_1\times\bar{P}_2\times\cdots}$$

がなりたつことを証明せよ. ($\overline{P_1\times P_2\times\cdots}$ を $P_1, P_2, \cdots$ の**完備直積**という.)

(iii) $\mathcal{B}_1, \mathcal{B}_2$ がそれぞれ Ω_1, Ω_2 の上の σ 加法族であるとき, $B\in\mathcal{B}_1\times\mathcal{B}_2$ の $\omega_1\in\Omega_1$ による切口は $\mathcal{B}_2$ に属することを示せ.

[ヒント] この性質をもつ B の全体がすべての $B_1\times B_2$ ($B_i\in\mathcal{B}_i$, $i=1,2$) を含む σ 加法族であることを示せばよい.

§2.4 標準確率空間

(Ω_1, P_1), (Ω_2, P_2) を確率空間とする. Ω_1 から Ω_2 への全単射, すなわち Ω_1 から Ω_2 の上への 1 対 1 写像 f が

$$A_1\in\mathcal{D}(P_1)\Leftrightarrow f(A_1)\in\mathcal{D}(P_2)\Longrightarrow P_1(A_1)=P_2(f(A_1))$$

をみたすとき, この写像 f を (Ω_1, P_1) から (Ω_2, P_2) への**強同型写像**という. このような f が存在するとき, (Ω_1, P_1) は (Ω_2, P_2) と**強同型**といい

$$(\Omega_1, P_1) \approx (\Omega_2, P_2)$$

であらわす．また強同型写像 f を指定したいときには，(Ω_1, P_1) は f によって (Ω_2, P_2) に強同型といい，

$$(\Omega_1, P_1) \approx (\Omega_2, P_2) \quad (f)$$

とかく．明らかに

反射律：$(\Omega, P) \approx (\Omega, P) \ (I)$，　I は恒等写像，

対称律：$(\Omega_1, P_1) \approx (\Omega_2, P_2) \ (f) \Longrightarrow (\Omega_2, P_2) \approx (\Omega_1, P_1) \ (f^{-1})$，

推移律：$(\Omega_1, P_1) \approx (\Omega_2, P_2) \ (f)$，　$(\Omega_2, P_2) \approx (\Omega_3, P_3) \ (g)$

$$\Longrightarrow (\Omega_1, P_1) \approx (\Omega_3, P_3) \ (g \circ f)$$

がなりたつから強同型関係は同値関係の一種である．

強同型よりもやや弱い概念として同型がある．確率空間 (Ω, P) の P 測度1の部分集合 Ω' をとり，P の Ω' への制限 $P'=P|_{\Omega'}$ を

$$\mathcal{D}(P') = \{A' \in \mathcal{D}(P) \,|\, A' \subset \Omega'\},$$

$$P'(A') = P(A'), \qquad A' \in \mathcal{D}(P')$$

で定義すると，P' は明らかに Ω' の上の確率測度である．確率空間 (Ω', P') を (Ω, P) の Ω' への制限という．(Ω_1, P_1) が (Ω_2, P_2) に同型であるというのは，(Ω_i, P_i) の P_i 測度1の部分集合 Ω_i' $(i=1,2)$ が存在して，(Ω_1, P_1) の Ω_1' への制限が (Ω_2, P_2) の Ω_2' への制限に強同型なことである．このとき

$$(\Omega_1, P_1) \sim (\Omega_2, P_2)$$

であらわす．同型も同値関係の一種である．反射律，対称律は明らかであるから，推移律だけを証明しておく．

$$(\Omega_1, P_1) \sim (\Omega_2, P_2), \qquad (\Omega_2, P_2) \sim (\Omega_3, P_3)$$

とすると，(Ω_1, P_1) の制限 (Ω_1', P_1')，(Ω_2, P_2) の制限 (Ω_2', P_2')，(Ω_2'', P_2'')，および (Ω_3, P_3) の制限 (Ω_3', P_3') があって

$$(\Omega_1', P_1') \approx (\Omega_2', P_2') \quad (f), \qquad (\Omega_2'', P_2'') \approx (\Omega_3', P_3') \quad (g)$$

となる．

$$\tilde{\Omega}_2 = \Omega_2' \cap \Omega_2'', \qquad \tilde{\Omega}_1 = f^{-1}(\tilde{\Omega}_2), \qquad \tilde{\Omega}_3 = g(\tilde{\Omega}_2)$$

とおくと，$P_i(\tilde{\Omega}_i)=1$ $(i=1, 2, 3)$ である．P_i の $\tilde{\Omega}_i$ への制限を $\tilde{P}_i$ とおくと，$(\tilde{\Omega}_i, \tilde{P}_i)$ は (Ω_i, P_i) の制限で，しかも

$$(\tilde{\Omega}_1, \tilde{P}_1) \approx (\tilde{\Omega}_2, \tilde{P}_2) \approx (\tilde{\Omega}_3, \tilde{P}_3)$$

となるから，$(\Omega_1, P_1) \sim (\Omega_3, P_3)$ が得られる．

定義 2.1 P を Ω の上の完備確率測度とする．(Ω, P) が $(\boldsymbol{R}^1, \mu)$ (μ は $\boldsymbol{R}^1$ の上の正則確率測度）に同型であるとき，(Ω, P) を**標準確率空間**，P を Ω の上の**標準確率測度**という．——

下に示すように，応用上有用な確率空間はほとんどすべて標準確率空間である．

定理 2.5 Ω が可算（有限または可算無限）のときには，Ω に定義域 2^{Ω} の確率測度 P をそえて得られる確率空間 (Ω, P) は標準確率空間である．

証明 Ω が可算無限のときを証明する．Ω の元を $\omega_1, \omega_2, \cdots$ とする．$\boldsymbol{R}^1$ の上に正則確率測度 Q を

$$Q(E) = \sum_{n \in E} P\{\omega_n\}$$

で定義すると，$Q(\boldsymbol{N})=1$ となる．$\boldsymbol{N}$ は自然数の集合である．Q の $\boldsymbol{N}$ への制限を Q' とすると，対応 $n \to \omega_n$ により，$(\boldsymbol{N}, Q')$ は (Ω, P) に強同型であるから，$(\boldsymbol{R}, Q)$ は (Ω, P) に同型となる．したがって (Ω, P) は標準である．∎

$\boldsymbol{R}^1$ の上の正則確率測度は明らかに標準であるが，$\boldsymbol{R}^n$ $(n=2, 3, \cdots)$ の上の正則確率空間もすべて標準である．もっと一般に完備可分距離空間の上の正則確率測度も標準であるが，このことを証明するために準備をする．

補題 2.1 P を完備可分距離空間 S の上の正則確率測度とする．任意の $A \in \mathcal{D}(P)$ と任意の $\varepsilon>0$ に対し，適当なコンパクト集合 $K \subset A$ をとって

$$P(A-K) < \varepsilon$$

とすることができる．

注意 P を位相空間 S の上の確率測度とし，$\mathcal{B}(S) \subset \mathcal{D}(P)$ とするとき，上の性質をもつ A を $\boldsymbol{K}$ **正則**といい，すべての P 可測集合が $\boldsymbol{K}$ 正則となるような P を $\boldsymbol{K}$ **正則**という．$\boldsymbol{K}$ 正則ならば，勿論正則である．この用語によれば，上の補題は次のようにいい表わされる．

"完備可分距離空間の上の正則確率測度は K 正則である."

補題 2.1 の証明 ρ を S の上の距離とする．始めに $A=S$ のときを考える．S が可分であるから，S の中で稠密な可算集合 $\{a_1, a_2, \cdots\}$ がある．任意の n に対し

$$S = \bigcup_{n=1}^{\infty} B_{nk}, \qquad B_{nk} = \left\{x \in S \,\middle|\, \rho(x, a_n) \leq \frac{1}{k}\right\} \quad \left(\text{中心 } a_n,\ \text{半径 } \frac{1}{k} \text{ の閉球}\right)$$

$$\bigcup_{n=1}^{N} B_{nk} \uparrow S \qquad (N \to \infty)$$

であるから，おのおのの k に対し，$N(k)$ を十分大きくとって，

$$P(S-B_k) < 2^{-(k+1)}\varepsilon, \qquad B_k = \bigcup_{n=1}^{N(k)} B_{nk}$$

とすることができる．B_k はすべて閉集合であるから，

$$K = \bigcap_{k=1}^{\infty} B_k$$

も閉集合である．またすべての k に対して

$$K \subset B_k = \bigcup_{n=1}^{N(k)} B_{nk}$$

であり，B_{nk} は直径 $2/k$ 以下であるから，K は全有界である．S が完備であるから，全有界閉部分集合 K はコンパクトである．しかも

$$P(S-K) = P\Big(S-\bigcap_k B_k\Big) = P\Big(\bigcup_k (S-B_k)\Big)$$
$$\leqq \sum_k P(S-B_k) < \varepsilon.$$

これで S が K 正則であることがわかった．

次に A を一般の Borel 集合とする．このとき任意の $\varepsilon>0$ に対して，閉集合 $F\subset A$ と開集合 $G\supset A$ をとって，

$$P(G-F) < \varepsilon$$

とできることをまず示そう．そのためにはこの性質をもつ A の全体 $\mathcal{B}$ が Borel 集合族 $\mathcal{B}(S)$ を含むことをいえばよい．距離空間では任意の開集合は閉集合の可算和であるから，$\mathcal{B}$ がすべての開集合を含むことは明らかである．$\mathcal{B}$ が補演算で閉じていることは，$G\supset A\supset F$ から $G^c\subset A^c\subset F^c$ がでることから明らかである．$A_n \in \mathcal{B}$ $(n=1, 2, \cdots)$ とし，

$$F_n \subset A_n \subset G_n, \qquad P(G_n-F_n) < 2^{-n}\varepsilon$$

なる閉集合列 $\{F_n\}$ と開集合列 $\{G_n\}$ をとると，

$$\bigcup_n F_n \subset \bigcup_n A_n \subset \bigcup_n G_n,$$

$$P\Big(\bigcup_n G_n - \bigcup_n F_n\Big) \leqq \sum_n P(G_n-F_n) \qquad (\text{例題 } 2.1\,(\text{iv}))$$
$$< \varepsilon$$

となる．したがって N を十分大きくとると，

$$P\Big(\bigcup_n G_n - \bigcup_{n=1}^{N} F_n\Big) < \varepsilon, \qquad \bigcup_{n=1}^{N} F_n \subset \bigcup_n A_n \subset \bigcup_n G_n$$

を得るが，これは $\bigcup_n A_n \in \mathcal{B}$ を意味する．したがって $\mathcal{B}$ は可算和でも閉じている．これから $\mathcal{B}$ は開集合族を含む σ 加法族であることがわかる．故に $\mathcal{B}\supset\mathcal{B}(S)$.

上のことから，任意の Borel 集合 A と任意の $\varepsilon>0$ に対して，閉集合 $F\subset A$ があって，

$$P(A-F) < \varepsilon.$$

P が正則であることから，これは任意の P 可測集合に対してもなりたつ．始めに S に対して導入したコンパクト集合 K をとると，$F\cap K$ は A のコンパクト部分集合で，

$$P(A-F\cap K) \leqq P(A-F)+P(A\setminus K) < 2\varepsilon.$$

ε は任意の正数であるから，これは A が K 正則であることを示している．したがって P は K 正則である．▌

補題 2.2 (Lusin の定理) S を完備可分な距離空間，T を可分距離空間とし，P を S の上の正則確率測度，$f: S\to T$ を P 可測（すなわち可測 $\mathcal{D}(P)/\mathcal{B}(T)$）とする．このとき任意の P 可測集合 $A\subset S$ と任意の $\varepsilon>0$ に対して適当なコンパクト集合 $K=K(A, \varepsilon)\subset A$ をとって，

$$P(A-K) < \varepsilon, \qquad \text{制限写像 } f_K \equiv f|_K : K \to T \text{ は連続}$$

となるようにできる．

証明 $S=T=\boldsymbol{R}^1$ の場合は Lusin の定理としてよく知られている．証明は本質的にこの特別の場合と変らない．S, T の上の距離を ρ, d であらわそう．T が可分であるから，稠密な可算集合 $\{b_1, b_2, \cdots\}\subset T$ がある．任意の k に対して

$$T = \bigcup_n B_{nk}, \qquad B_{nk} = \left\{y \in T \,\middle|\, d(y, b_n) \leqq \frac{1}{k}\right\}.$$

さて

$$A_{nk} = A\cap f^{-1}(B_{nk}), \qquad A_{nk}' = A_{nk} \setminus \bigcup_{i<n} A_{ik}$$

とおくと，これは P 可測である．しかも

$$A = A\cap S = A\cap f^{-1}(T)$$
$$= \bigcup_n A\cap f^{-1}(B_{nk}) = \bigcup_n A_{nk} = \sum_n A_{nk}'.$$

前の補題により，コンパクト部分集合 $K_{nk}\subset A_{nk}'$ があって

$$P(A_{nk}'-K_{nk})\leqq 2^{-n-k}\varepsilon,$$

したがって

$$P\Big(A-\sum_n K_{nk}\Big)\leqq \sum_n P(A_{nk}-K_{nk}) \qquad (\text{例題 2.1 (iv)})$$
$$<2^{-k}\varepsilon.$$

$N(k)$ を十分大きくとって，$K_k=\sum_{n=1}^{N(k)}K_{nk}$ とおくと，K_k はコンパクトで，

$$P(A-K_k)<2^{-k}\varepsilon.$$

さらに $K=\bigcap_{k=1}^{\infty}K_k$ とおくと，K もコンパクトである．しかも

$$P(A-K)\leqq \sum_k P(A-K_k)<\varepsilon.$$

$f_K=f|_K:K\to T$ の連続性をいえば，補題の証明は終る．まず $K\subset K_k$ から

$$K=K\cap K_k=\sum_{n=1}^{N(k)}K\cap K_{nk}$$

がでる．$K\cap K_{nk}$ の中に空なものがあれば，この和からはずしておく．K_{nk} の中に a_{nk} を任意にとり，これを固定して，$g_k:K\to T$ を

$$g_k(x)=f(a_{nk}), \qquad x\in K_{nk}, \qquad n=1,2,\cdots,N(k)$$

と定義する．$K\cap K_{nk}$ $(n=1,2,\cdots,N(k))$ は互いに素なコンパクト集合で，g_k は各集合の上で一定であるから，g_k は連続である．また $x\in K\cap K_{nk}$ に対し，$f(a_{nk}),f(x)\in f(K_{nk})\subset f(A_{nk})\subset B_{nk}$ であるから，

$$d(g_k(x),f(x))=d(f(a_{nk}),f(x))\leqq \frac{2}{k}, \qquad n=1,2,\cdots,N(k).$$

したがって $x\in K$ に対し

$$d(g_k(x),f_K(x))<\frac{2}{k}$$

となり $\{g_k\}$ は f_K に一様収束する．故に f_K は連続である．∎

S,T を一般の集合とし，P を S の上の確率測度，$f:S\to T$ を任意の写像とする．このとき T の上に集合関数 Q を

$$\mathcal{D}(Q)=\{B\subset T\mid f^{-1}(B)\in\mathcal{D}(P)\}, \qquad Q(B)=P(f^{-1}(B))$$

で定義する．Q が T の上の確率測度であることは明らかである．この Q を P の f による**像測度**といい，Pf^{-1} または fP であらわす．P が完備ならば，$Q=Pf^{-1}$ も完備である．

補題 2.3 S を完備可分距離空間，T を可分距離空間，P を S の上の正則確率測度，f を S から T の中への P 可測写像とする．このとき像測度 $Q=Pf^{-1}$ は T の上の K 正則確率測度である．

証明 Q が T の上の完備確率測度であること，$\mathcal{D}(Q)$ が $\mathcal{B}(T)$ を含むことは明らかである．したがって Q が正則であることをいうには，"$B\in\mathcal{D}(Q)$ と $\varepsilon>0$ に対し適当なコンパクト集合 $K\subset B$ をとって，$Q(B-K)<\varepsilon$ とできる" ことをいえばよい．$B\in\mathcal{D}(Q)$ とすれば，$A=f^{-1}(B)$ が P 可測であるから，前補題により，適当なコンパクト集合 $H\subset A$ をとって

$$P(A-H)<\varepsilon, \qquad f_H=f|_H:H\to T \text{ は連続}$$

とすることができる．したがって $K=f_H(H)=f(H)$ はコンパクトである．明らかに $f^{-1}(K)\supset H$ であるから，

$$Q(B-K)=P(f^{-1}(B-K))=P(f^{-1}(B)-f^{-1}(K))\leqq P(A-H)<\varepsilon. \quad ∎$$

これだけの準備をもとにして，目的の次の定理を証明する．

定理 2.6 完備可分距離空間 S の上の正則確率測度 P は標準である．

証明 補題 2.1 で用いた集合

$$B_{nk}, \qquad n,k=1,2,\cdots$$

をならべて $B_1,B_2,\cdots$ とし，B_n の指示関数を e_n であらわす．さて S の上の実関数 $f:S\to\boldsymbol{R}^1$ を

$$f(x)=\sum_{n=1}^{\infty}\frac{2e_n(x)}{3^n}$$

で定義する．$f(S)$ が Cantor 集合 $\boldsymbol{K}$ の部分集合であり，f が P 可測であることは，測度論の基礎知識があれば，容易に証明できるから，認められたものとする．$Q=Pf^{-1}$ とすると，$\boldsymbol{R}^1$ が可分距離空間であるから，前補題により，Q は $\boldsymbol{R}^1$ の上の正則確率測度である．

残るのは (S,P) が $(\boldsymbol{R}^1,Q)$ に同型なことの証明である．x,y が S の異なる点とすれば，適当な n に対して $x\in B_n$，$y\notin B_n$ となるから，$e_n(x)\neq e_n(y)$，したがって $f(x)\neq f(y)$．故に f は単射であり，f の値域を $T=f(S)\,(\subset\boldsymbol{R}^1)$ に制限して写像 $f':S\to T$ を定めると，f' は全単射である．$f'^{-1}(T)=S$ であるから，$Q(T)=1$

であって，$Q'=Q|_T$ は T の上の確率測度である．$f(A)=f'(A)\subset T$ であるから，

$$A\in\mathcal{D}(P)\iff f'(A)\in\mathcal{D}(Q')\implies Q'(f'(A))=P(A)$$

となることがすぐにわかる．これから

$$(S,P)\approx(T,Q')\quad(f')$$

がでて，結局 (S,P) と $(\boldsymbol{R}^1,Q)$ とは同型となる．▌

可算集合 Ω は距離

$$\rho(x,y)=1\quad(x\neq y),\qquad =0\quad(x=y)$$

により完備可分であり，このとき $\mathcal{B}(\Omega)=2^\Omega$ となるから，定理2.5は定理2.6に含まれていると考えてよい．

可算個の距離空間 $S_n=(S_n,\rho_n)$ (ρ_n は距離，$n=1,2,\cdots$) の積空間 $S=\prod_n S_n$ の位相（積位相）は距離

$$\rho((x_n),(y_n))=\sum_{n=1}^\infty 2^{-n}[\rho_n(x_n,y_n)\wedge 1],\qquad a\wedge b=\min(a,b)$$

で与えられるが，$S_n=(S_n,\rho_n)$ が完備可分ならば，$S=(S,\rho)$ も完備可分である．これにより

$$\boldsymbol{R}^\infty=\boldsymbol{R}^1\times\boldsymbol{R}^1\times\cdots$$

も完備可分距離空間であることがわかる．$\boldsymbol{R}^\infty$ は実数列全体の空間であるから，**数列空間**とよばれることもある．上の定理により，数列空間 $\boldsymbol{R}^\infty$ の上の正則確率測度も標準である．

$\boldsymbol{R}^\infty$ の Borel 集合族 $\mathcal{B}(\boldsymbol{R}^\infty)$ を $\mathcal{B}^\infty$ とかく．さて

$$\mathcal{B}^\infty=\mathcal{B}^1\times\mathcal{B}^1\times\cdots$$

を証明しておこう．n 射影 $\pi_n(x_1,x_2,\cdots)=x_n$ は連続であるから，$\pi_n^{-1}(E)$ はすべての n，すべての $E\in\mathcal{B}^1$ に対して，$\mathcal{B}^\infty$ に属する．これから

$$\mathcal{B}^\infty\supset\mathcal{B}^1\times\mathcal{B}^1\times\cdots.$$

逆の包含関係を示すには，$\boldsymbol{R}^\infty$ の（積位相に関する）開集合が

$$(a_1,b_1)\times(a_2,b_2)\times\cdots\times(a_n,b_n)\times\boldsymbol{R}^1\times\boldsymbol{R}^1\times\cdots$$

の形の集合の可算和であらわされることに注意すればよい．この事情は可分距離空間（一般に可算開基をもつ位相空間）の可算積についても同様である．

$$(\Omega_n,P_n)\approx(\Omega_n',P_n')\quad(f_n),\qquad n=1,2,\cdots$$

ならば，

$$\left(\prod_n\Omega_n,\prod_n P_n\right)\approx\left(\prod_n\Omega_n',\prod_n P_n'\right)\quad(f),$$

ここに f は

$$f(\omega_1,\omega_2,\cdots)=(f_1(\omega_1),f_2(\omega_2),\cdots)$$

で定義する．

また $\Omega=\prod_n\Omega_n$，$P=\prod_n P_n$ とおき，$P_n(A_n)=1$ とすると

$$\begin{aligned}P\left(\prod_n A_n\right)&=P\left(\lim_{n\to\infty}A_1\times A_2\times\cdots\times A_n\times\Omega_{n+1}\times\Omega_{n+2}\times\cdots\right)\\&=\lim_{n\to\infty}P(A_1\times A_2\times\cdots\times A_n\times\Omega_{n+1}\times\Omega_{n+2}\times\cdots)\\&=\lim_{n\to\infty}P_1(A_1)P_2(A_2)\cdots P_n(A_n)=1\end{aligned}$$

となる．

以上の二つの事実を用いて

$$(\Omega_n,P_n)\sim(\Omega_n',P_n')\ (n=1,2,\cdots)\implies\left(\prod_n\Omega_n,\prod_n P_n\right)\sim\left(\prod_n\Omega_n',\prod_n P_n'\right)$$

が容易に証明される．

また

$$(\Omega,P)\sim(\Omega',P')\implies(\Omega,\bar{P})\sim(\Omega',\bar{P}')\qquad(\bar{\ }\text{ は Lebesgue 拡大})$$

もほとんど明らかである．

さて μ_n を $\boldsymbol{R}^1$ の上の Borel 確率測度とすれば直積測度 $\mu=\prod_{n=1}^\infty\mu_n$ の定義域は $\mathcal{B}^1\times\mathcal{B}^1\times\cdots$ であるが，これは上に注意したように $\mathcal{B}^\infty=\mathcal{B}(\boldsymbol{R}^\infty)$ に等しいから，μ は $\boldsymbol{R}^\infty$ の上の Borel 確率測度となる．

例題2.3(ii)により

$$\overline{\prod_{n=1}^\infty\mu_n}=\overline{\prod_{n=1}^\infty\bar{\mu}_n}.$$

これらの事実から，次の定理が証明される．

定理 2.7　P_n を完備可分距離空間 S_n の上の正則確率測度とすれば，P_n ($n=1,2,\cdots$) の完備直積 $P=\overline{\prod_n P_n}$ は $S=\prod_n S_n$（これが完備可分距離空間であることはすでにのべた）の上の正則確率測度（したがって標準確率測度）である．

定理 2.8　標準確率測度 P_n ($n=1,2,\cdots$) の完備直積 $P=\overline{\prod_n P_n}$ も標準確率測度である．

例題 2.4 T を完備可分距離空間 S の Borel 部分集合とする．S の中の距離 ρ を T に制限したものを ρ_T とすれば，T は ρ_T に関して可分距離空間となる．T の上の正則確率空間 (T,P) は標準であることを示せ．

［ヒント］ $Q(B)=P(B\cap T)$ とおくと，$T\in\mathcal{B}(S)$ と P の正則性により，Q も T の上の正則確率測度である．したがって (S,Q) は標準である．$T\in\mathcal{B}(S)$ により，$T\in\mathcal{D}(Q)$ かつ $Q(T)=1$ が得られる．したがって P と $Q|_T$ とは一致し，$(T,P)\sim(S,Q)$ がわかる．

§2.5 1次元の分布

μ を $\boldsymbol{R}^1$ の上の正則確率測度とすれば，$\mu|_{\mathcal{B}^1}$ は $\boldsymbol{R}^1$ の上の Borel 確率測度であり，μ は $\mu|_{\mathcal{B}^1}$ の Lebesgue 拡大であるから，$\mu(E)$ $(E\in\mathcal{B}^1)$ を与えると，μ は完全に定まる．μ はまたその分布関数(§2.2)

$$F(x)=\mu(-\infty,x]$$

によっても完全に決定される．

$\boldsymbol{R}^1$ の上の正則確率測度を簡単に**1次元の分布**または**分布**ということにする．

a) Lebesgue の分解

分布を $\mu,\nu,\cdots$ であらわす．

$$\mu\{a\}>0$$

のとき，a を μ の**不連続点**という．これは μ の分布関数の飛躍点である．μ の不連続点の集合 D_μ は可算である．これに対し

$$\mu\{a\}=0$$

となる a は μ の**連続点**とよばれる．μ の連続点の集合を C_μ であらわそう．

$\mu(D_\mu)=1$ となるような μ を**純不連続分布**という．このときには，$D_\mu=\{a_n\,|\,n=1,2,\cdots\}$ とすると，任意の集合 E に対して

$$\mu(E)=\sum_{a_n\in E}\mu\{a_n\}$$

となるから，μ は $\{a_n\}$ と $\{p_n=\mu\{a_n\}\}$ で完全に定まる．したがってこのような μ は

$$\begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix} \quad \text{または} \quad \begin{pmatrix} a_n \\ p_n \end{pmatrix}_{n=1,2,\cdots}$$

であらわす．明らかに

$$p_n\geqq 0,\qquad \sum_n p_n=1$$

である．確率論でしばしばあらわれる純不連続分布の例をあげておこう．

δ 分布 $\quad\delta_a=\begin{pmatrix} a \\ 1 \end{pmatrix}$ すなわち $\delta_a(E)=\begin{cases} 1, & a\in E, \\ 0, & a\notin E. \end{cases}$

特に $a=0$ のときには単に δ であらわす．

2項分布 $\quad b_{n,p}=\begin{pmatrix} k \\ \binom{n}{k}p^k(1-p)^{n-k} \end{pmatrix}_{k=0,1,2,\cdots,n}$,

ここで n,p はパラメータで $n=1,2,\cdots$，$0<p<1$．

Poisson 分布 $\quad p_\lambda=\begin{pmatrix} k \\ e^{-\lambda}\lambda^k/k! \end{pmatrix}_{k=0,1,2,\cdots}$.

すべての点が μ の連続点であるとき，μ を**連続分布**という．μ が連続ならば，すべての a に対し $\mu\{a\}=0$，したがってすべての可算集合 C に対して $\mu(C)=0$ である．連続より強い条件として絶対連続がある．E の Lebesgue 測度が 0 であれば，$\mu(E)$ が必ず 0 であるとき，μ を**絶対連続**という．測度論の Radon-Nikodym の定理によれば絶対連続であるための必要十分条件は密度をもつことである．したがって絶対連続分布は密度関数 $f(x)$ を用いて

$$\mu(E)=\int_E f(x)dx\quad(E\in\mathcal{B}^1)\qquad\text{または}\qquad \mu(dx)=f(x)dx$$

であらわされる．ここに f は Lebesgue 可測で

$$f(x)\geqq 0,\qquad \int_{R^1}f(x)dx=1$$

をみたす．Gauss 分布，Cauchy 分布(例題 2.1 (i))は絶対連続である．

連続でも絶対連続とは限らない．実際，ある E に対して

$$\mu(E)=1,\qquad \lambda(E)=0\quad(\lambda \text{ は Lebesgue 測度})$$

となる連続分布 μ もある．このような分布 μ は**特異分布**とよばれる．特異分布の例をあげよう．P を $[0,1)$ の上の Lebesgue 測度とし，$f:[0,1)\to\boldsymbol{R}^1$ を

$$x=\sum_{n=1}^{\infty}\frac{\varepsilon_n}{2^n}\ \ (2\text{ 進法展開})\longrightarrow f(x)=\sum_{n=1}^{\infty}\frac{2\varepsilon_n}{3^n}$$

で定める．$k/2^n$ の形の x には末尾が $000\cdots$ のものと，$111\cdots$ のものの二通りの2進法展開があるが，前者をとることにする．こうすれば，ε_n は x の関数として

$$\varepsilon_n(x) = [2^n x] - 2[2^{n-1}x] \qquad ([\ ]\text{は整数部分})$$

であらわされる．これは x の Borel 可測関数であるから，$f(x)$ も Borel 可測（したがって P 可測）である．故に補題 2.3 により，像測度 fP は $\boldsymbol{R}^1$ の上の正則確率測度である．これを μ とする．明らかに $f([0,1])$ は Cantor 集合 $\boldsymbol{K}$ に含まれ，かつ f は狭義増加関数であるから，

$$\mu(\boldsymbol{K}) = 1 \quad \text{かつ} \quad \mu\{a\} = 0 \qquad (a \in \boldsymbol{R}^1)$$

であるが，$\boldsymbol{K}$ の Lebesgue 測度は 0 であるから，μ は特異分布である．

$\mu_1, \mu_2, \cdots$ が分布ならば

$$\mu(E) = \sum_n c_n \mu_n(E), \qquad E \in \mathcal{B}^1 \qquad (\text{ただし } c_n \geqq 0,\ \sum_n c_n = 1)$$

も分布である．これを $\{\mu_n\}$ の**凸結合**という．

定理 2.9（Lebesgue の分解定理）　任意の分布は純不連続分布，絶対連続分布，特異分布の凸結合であらわされる．（この3種の分布が全部あらわれるとは限らない．）しかもこの分解の方法は一通りである．

証明　μ を分布とし，任意の Borel 集合 E に対し

$$\nu_d(E) = \mu(E \cap D_\mu), \qquad \nu_c(E) = \mu(E \cap C_\mu)$$

$$\mu = \nu_d(\boldsymbol{R}^1)\frac{\nu_d}{\nu_d(\boldsymbol{R}^1)} + \nu_c(\boldsymbol{R}^1)\frac{\nu_c}{\nu_c(\boldsymbol{R}^1)}$$

（係数が 0 のときには，その項をとり去る）

とおくと，μ が純不連続分布と連続分布との凸結合になる．したがって連続分布が特異分布と絶対連続分布との凸結合になることをいえばよい．μ を連続とし，

$$s = \sup\{\mu(E) \mid E \in \mathcal{B}^1, \lambda(E) = 0\}$$

とおくと，集合列 $E_n \in \mathcal{B}^1$ が存在して

$$\lambda(E_n) = 0, \qquad \mu(E_n) \longrightarrow s$$

となる．$S = \bigcup_n E_n$ とおくと

$$\lambda(S) = 0, \qquad \mu(S) = s,$$

したがって

$$A \subset S^c,\ \lambda(A) = 0 \Longrightarrow \mu(A) = 0$$

となる．このことを頭において

$$\nu_s(E) = \mu(E \cap S), \qquad \nu_{ac}(E) = \mu(E \cap S^c), \qquad E \in \mathcal{B}^1$$

とおくと，

$$\mu = \nu_s(\boldsymbol{R}^1)\frac{\nu_s}{\nu_s(\boldsymbol{R}^1)} + \nu_{ac}(\boldsymbol{R}^1)\frac{\nu_{ac}}{\nu_{ac}(\boldsymbol{R}^1)}$$

（係数が 0 のときには，その項をとり去る）

が求める分解であることがわかる．▌

b）分布列の収束

分布列 μ_n が分布 μ に収束するとは，任意の有界連続実関数 f に対して

$$\int_{R^1} f(x)\mu_n(dx) \longrightarrow \int_{R^1} f(x)\mu(dx) \qquad (n \to \infty)$$

となることであると定義する．

すべての $E \in \mathcal{B}^1$ に対し

$$\mu_n(E) \longrightarrow \mu(E) \qquad (n \to \infty)$$

ならば，$\mu_n \to \mu$ となることは容易に証明できるが，逆は成立しないことは次の諸例から明らかである．

$a_n \to a$ ならば，$\delta_{a_n} \to \delta_a$．しかし $E = \{a\}$ に対しては

$$\delta_{a_n}(E) = 0 \quad (a_n \neq a), \qquad \delta_a(E) = 1.$$

$v_n \to 0$ ならば，$N_{0,v_n} \to \delta$．なぜならば，有界連続な f に対し

$$\begin{aligned}\int_{R^1} f(x) N_{0,v_n}(dx) &= \int_{R^1} f(x)\frac{1}{\sqrt{2\pi v_n}}e^{-x^2/2v_n}dx \\ &= \int_{R^1} f(\sqrt{v_n}\,y)\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy \\ &\longrightarrow \int_{R^1} f(0)\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy \qquad (\text{有界収束定理}) \\ &= f(0) = \int_{R^1} f(x)\delta(dx).\end{aligned}$$

しかし $E = [0, \infty)$ に対しては

$$N_{0,v_n}(E) = \frac{1}{2}, \qquad \delta(E) = 1.$$

次の定理は分布の列の収束に関するいくつかの必要十分条件を与える．

定理 2.10　次の条件は互いに同等である．F_n, F はそれぞれ μ_n, μ の分布関数である．

(i) $\mu_n \longrightarrow \mu$,

(ii) コンパクトな台をもつ任意の連続関数 g に対し

$$\int_{R^1} g(x)\mu_n(dx) \longrightarrow \int_{R^1} g(x)\mu(dx)$$

($\{x \mid f(x) \neq 0\}$ の閉包を f の**台**という),

(iii) $\mu(E^o)=\mu(\bar{E})$ (E^o, $\bar{E}$ はそれぞれ E の開核, 閉包) なるすべての E に対し

$$\mu_n(E) \longrightarrow \mu(E),$$

(iv) F のすべての連続点 x に対し

$$F_n(x) \longrightarrow F(x),$$

(v) $\boldsymbol{R}^1$ で稠密な可算集合 C があって, C のすべての点 x に対し

$$F_n(x) \longrightarrow F(x).$$

証明 (iii) $\Longrightarrow$ (iv) $\Longrightarrow$ (v) は明らか. したがって

$$(v) \Longrightarrow (ii) \Longrightarrow (i) \Longrightarrow (iii)$$

をいえばよい.

(v) を仮定する. g がコンパクトな台をもつ連続関数とすると, g は一様連続. したがってコンパクトな台をもつ左連続階段関数 g_ε があって

$$\sup_x |g_\varepsilon(x)-g(x)| < \varepsilon$$

となる. しかも, 仮定 (v) の C は稠密であるから, g_ε の飛躍点 $a_i=a_i(\varepsilon)$ ($i=0, 1, 2, \cdots, m=m(\varepsilon)$) はすべて C に属すると仮定してよい. 簡単のため積分

$$\int g(x)\nu(dx)$$

を (g, ν) とかくと,

$$\begin{aligned}|(g, \mu_n)-(g, \mu)| &\leqq |(g, \mu_n)-(g_\varepsilon, \mu_n)|+|(g, \mu)-(g_\varepsilon, \mu)| \\ &\quad +|(g_\varepsilon, \mu_n)-(g_\varepsilon, \mu)| \\ &\leqq \varepsilon+\varepsilon+\left|\sum_{i=1}^m g_\varepsilon(a_i)(F_n(a_i)-F_n(a_{i-1})) \right. \\ &\quad \left. -\sum_{i=1}^m g_\varepsilon(a_i)(F(a_i)-F(a_{i-1}))\right|.\end{aligned}$$

仮定 (v) により, すべての $i=0, 1, 2, \cdots, m$ に対し

$$\lim_{n\to\infty} F_n(a_i) = F(a_i)$$

であるから, 上の最後の式の第3項は $n\to\infty$ のとき0に収束する. したがって

$$\limsup_{n\to\infty} |(g, \mu_n)-(g, \mu)| \leqq 2\varepsilon.$$

ε は任意の正数であるから, 上の左辺は0. これで (v) $\Longrightarrow$ (ii) が証明された.

(ii) を仮定する. f を任意の有界連続関数とせよ. 任意の $m>0$ に対し, $[-m, m]$ の上で f と一致し, $[-m-1, m+1]$ の外では0となる連続関数 g が存在する. さて明らかに

$$|(f, \mu_n)-(f, \mu)| \leqq (|f-g|, \mu_n)+(|f-g|, \mu)+|(g, \mu_n)-(g, \mu)|.$$

$f-g$ は有界であるから,

$$|f(x)-g(x)| \leqq c$$

をみたす (x に無関係な) 定数 c をとることができる. 今

$$\begin{array}{lll} [-m+1, m-1] & \text{の上で} & h(x)=0, \\ [-m, m] & \text{の外で} & h(x)=c, \\ \text{これ以外の点で} & & 0 \leqq h(x) \leqq c \end{array}$$

となる連続関数 $h(x)$ をとると, $f-g$ が $[-m, m]$ 上で0であるから,

$$|f(x)-g(x)| \leqq h(x).$$

したがって

$$(|f-g|, \mu_n) \leqq (h, \mu_n) = (c, \mu_n)-(c-h, \mu_n) = c-(c-h, \mu_n).$$

$c-h$ は $[-m, m]$ の外では0に等しいから, (ii) の仮定により, $n\to\infty$ のとき

$$(c-h, \mu_n) \longrightarrow (c-h, \mu) = c-(h, \mu),$$

したがって

$$\limsup_{n\to\infty} (|f-g|, \mu_n) \leqq (h, \mu) \leqq c\mu([-m+1, m-1]^c),$$

また

$$(|f-g|, \mu) \leqq (h, \mu) \leqq c\mu([-m+1, m-1]^c).$$

仮定により

$$\lim_{n\to\infty} |(g, \mu_n)-(g, \mu)| = 0.$$

かくして

$$\limsup_{n\to\infty} |(f, \mu_n)-(f, \mu)| \leqq 2c\mu([-m+1, m-1]^c).$$

$m\to\infty$ とすると右辺は0に収束するから，左辺は0となる．これで(ii)$\Longrightarrow$(i)が証明された．

(i)を仮定する．E を(iii)の条件をみたす集合とする．R^1 の中では任意の開集合は増加閉集合列の極限であり，任意の閉集合は減少開集合列の極限であるから，任意の $\varepsilon>0$ に対し閉集合 $F_\varepsilon\subset E^0$ と開集合 $G_\varepsilon\supset\bar{E}$ があって

$$\mu(E^0)-\mu(F_\varepsilon)<\varepsilon,\quad \mu(G_\varepsilon)-\mu(\bar{E})<\varepsilon.$$

$\mu(E^0)=\mu(\bar{E})$ の仮定により，これは $\mu(E)$ に等しく

$$\mu(G_\varepsilon)<\mu(E)+\varepsilon,\quad \mu(F_\varepsilon)>\mu(E)-\varepsilon.$$

開集合 G_ε と閉集合 $\bar{E}$ に対応して連続関数 f_ε をとって

$$G_\varepsilon \text{ の外で } f_\varepsilon=0,\quad \bar{E} \text{ の上で } f_\varepsilon=1,\quad \text{いたる所で } 0\leqq f_\varepsilon\leqq 1$$

とすると，仮定(i)により

$$\lim_{n\to\infty}(f_\varepsilon,\mu_n)=(f_\varepsilon,\mu).$$

$$\mu_n(E)\leqq\mu_n(\bar{E})\leqq(f_\varepsilon,\mu_n)$$

により

$$\limsup_{n\to\infty}\mu_n(E)\leqq(f_\varepsilon,\mu)\leqq\mu(G_\varepsilon)\leqq\mu(E)+\varepsilon.$$

故に

$$\limsup_{n\to\infty}\mu_n(E)\leqq\mu(E).$$

同様に E^0 と F_ε とに対応して連続関数 g_ε をとって

$$\liminf_{n\to\infty}\mu_n(E)\geqq\mu(E)$$

が証明され，結局

$$\lim_{n\to\infty}\mu_n(E)=\mu(E)$$

が得られる．これで(i)$\Longrightarrow$(iii)が証明された．▌

上の定理から“$\mu_n\to\mu, \mu_n\to\nu\Longrightarrow\mu=\nu$”が得られる．実際 F, G をそれぞれ μ, ν の分布関数とすれば，$D_\mu\cup D_\nu$ の外では $F=G$. F, G はともに右連続で，$D_\mu\cup D_\nu$ は可算であるから，到る所 $F=G$, すなわち $\mu=\nu$. また $\{\mu_n\}$ が μ に収束すれば，$\{\mu_n\}$ のすべての部分列は μ に収束する．

定理 2.11 分布を元とする集合 $\mathcal{M}$ に対して，次の二つの条件は同等である．

(i) $\mathcal{M}$ の中の任意の無限列は収束部分列をもつ(極限分布は $\mathcal{M}$ の中にある必要はない)，

(ii) $\displaystyle\lim_{a\to\infty}\inf_{\mu\in\mathcal{M}}\mu[-a,a]=1.$

証明 (ii)がなりたたないとすると，適当な $\varepsilon_0>0$ と $\mathcal{M}$ の中の無限列 $\{\nu_n\}$ が存在して

$$\nu_n[-n,n]\leqq 1-\varepsilon_0,\quad n=1,2,\cdots.$$

もし(i)がなりたつならば，$\{\nu_n\}$ は収束部分列 $\nu_{n(k)}$ $(\to\nu)$ をもつ．$l\geqq k$ ならば，

$$\nu_{n(l)}[-n(k),n(k)]\leqq\nu_{n(l)}[-n(l),n(l)]\leqq 1-\varepsilon_0.$$

$(-n(k),-n(k)+1)$, $(n(k)-1,n(k))$ の中にそれぞれ ν の連続点 α_k, β_k をとると，$(\alpha_k,\beta_k]$ は $[-n(k),n(k)]$ に含まれるから，

$$\nu_{n(l)}(\alpha_k,\beta_k]\leqq\nu_{n(l)}[-n(k),n(k)]\leqq 1-\varepsilon_0,\quad l=k,k+1,\cdots.$$

$l\to\infty$ として，定理2.10により

$$\nu(\alpha_k,\beta_k]\leqq 1-\varepsilon_0.$$

$k\to\infty$ として

$$\nu(-\infty,\infty)\leqq 1-\varepsilon_0.$$

これは ν が分布であることに反する．これで(i)$\Longrightarrow$(ii)が証明された．

逆に(ii)がなりたつとし，$\{\mu_n\}$ を $\mathcal{M}$ の中の任意の部分列とする．R^1 の中に稠密な点列 $\{a_m\}$ をとる．μ_n の分布関数を F_n であらわすと，

$$0\leqq F_n(a_m)\leqq 1,\quad n,m=1,2,\cdots$$

により，a_m に対し $\{F_n\}$ の部分列 $\{F_{n(m,k)}, k=1,2,\cdots\}$ をとって

$$F_{n(1,k)}(a_1)\longrightarrow\tilde{F}(a_1)$$
$$F_{n(2,k)}(a_2)\longrightarrow\tilde{F}(a_2)$$
$$\cdots\cdots\cdots\cdots\cdots$$

とすることができる．しかも $\{F_{n(m+1,k)}, k=1,2,\cdots\}$ は $\{F_{n(m,k)}, k=1,2,\cdots\}$ の部分列となるようにとれる．$\{F_{n(k,k)}, k=1,2,\cdots\}$ の尾部は，任意の m に対し，$\{F_{n(m,k)}, k=1,2,\cdots\}$ の部分列であるから，すべての m に対し

$$F_{n(k,k)}(a_m)\longrightarrow\tilde{F}(a_m)\quad(\text{対角線論法!}).$$

明らかに

$$0\leqq\tilde{F}(a_m)\leqq 1,\quad \tilde{F}(a_m)\leqq\tilde{F}(a_l)\quad(a_m\leqq a_l).$$

仮定(ii)と $\{a_m\}$ の稠密性により，任意の ε に対し

$$F_n(a_m)-F_n(a_l)>1-\varepsilon,\quad n=1,2,\cdots$$

をみたす a_m, a_l がある．したがって

$$\tilde{F}(\infty)-\tilde{F}(-\infty) \geqq \tilde{F}(a_m)-\tilde{F}(a_l) \geqq 1-\varepsilon.$$

故に左辺は1に等しく，$\tilde{F}(\infty)=1$，$\tilde{F}(-\infty)=0$．さて

$$F(x) = \inf_{a_m>x} \tilde{F}(a_m)$$

とすれば，$F(x)$ が分布関数の条件をみたしていることが $\tilde{F}$ の性質からわかるから，これに対応ずる分布を ν とする．もし F の任意の連続点 a で

$$G_k(a) \equiv F_{n(k,k)}(a) \longrightarrow F(a) \qquad (k\to\infty)$$

が証明できれば，$\nu_{n(k,k)}\to\nu$ となり，証明は完了する．F の a における連続性により，任意の $\varepsilon>0$ に対し，$\delta=\delta(\varepsilon)>0$ をとって

$$F(a)-\varepsilon < F(a-\delta)$$

とすることができる．次に $\{a_m\}$ の稠密性により，

$$a-\delta < a_m < a$$

なる a_m がある．F の定義により

$$F(a)-\varepsilon < F(a-\delta) \leqq \tilde{F}(a_m) = \lim_{k\to\infty} G_k(a_m) \leqq \liminf_{k\to\infty} G_k(a).$$

また F の定義により，$a_l>a$ を適当にとって

$$F(a)+\varepsilon > \tilde{F}(a_l) = \lim_{k\to\infty} G_k(a_l) \geqq \limsup_{k\to\infty} G_k(a).$$

ε は任意の正数であるから，これから

$$F(a) \leqq \liminf_{k\to\infty} G_k(a), \qquad F(a) \geqq \limsup_{k\to\infty} G_k(a),$$

すなわち

$$F(a) = \lim_{k\to\infty} G_k(a).$$ ▮

c）分布の能率

分布 μ の p 次の能率 $M_p(\mu)$ $(p=1,2,\cdots)$ を

$$M_p(\mu) = \int_R x^p \mu(dx)$$

で定義する．$M_p(\mu)$ は上の積分が意味をもつときにのみ定義される．また

$$|M|_p(\mu) = \int_R |x|^p \mu(dx) \ (\in [0,\infty])$$

を μ の p 次の**絶対能率**という．$|M|_p(\mu)<\infty$ のときには $M_p(\mu)$ は有限確定である．絶対能率 $|M|_p(\mu)$ は，p が任意の正数のときにも定義される．$(|M|_p(\mu))^{1/p}$

は p とともに増大するが，これを証明するには次の補題が必要である．

補題 2.4（**Hölder の不等式**） $1<p,\ q<\infty,\ p^{-1}+q^{-1}=1$ のとき

$$\int_{R^1} |f(x)g(x)|\mu(dx) \leqq \left(\int_{R^1} |f(x)|^p \mu(dx)\right)^{1/p} \left(\int_{R^1} |g(x)|^q \mu(dx)\right)^{1/q}.$$

証明 右辺の二つの因数を a,b であらわす．a,b の一方が0または ∞ のときには上の不等式は明らかである．（便宜上 $0\cdot\infty=\infty\cdot 0=0$ とする．）したがって $0<a,\ b<\infty$ のときを考慮すればよい．f,g をそれぞれ $f/a, g/b$ でおきかえることにより，$a=b=1$ の場合に帰着される．

$$F(x) = -\log x, \qquad 0<x<\infty$$

とおくと

$$F''(x) = \frac{1}{x^2} > 0$$

であるから，$F(x)$ は凸関数（下に凸）である．したがって補題の条件をみたす p, q に対して

$$-\frac{1}{p}\log\alpha^p - \frac{1}{q}\log\beta^q \geqq -\log\left(\frac{1}{p}\alpha^p + \frac{1}{q}\beta^q\right), \qquad \alpha,\beta>0,$$

すなわち

$$\alpha\beta \leqq \frac{1}{p}\alpha^p + \frac{1}{q}\beta^q, \qquad \alpha,\beta>0.$$

この不等式は α,β が0のときにも当然なりたつ．故に

$$|f(x)g(x)| \leqq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$$

がすべての $x\in \boldsymbol{R}^1$ に対してなりたつ．これを積分すれば補題の証明が完了する．▮

定理 2.12 $1 \leqq p < q < \infty \Longrightarrow (|M|_p(\mu))^{1/p} \leqq (|M|_q(\mu))^{1/q}$.

証明 $r=q/p,\ s=q/(q-p)$ とおくと，

$$\frac{1}{r}+\frac{1}{s} = 1$$

となる．Hölder の不等式を用いて

$$\int_{R^1} |x|^p \mu(dx) \leqq \left(\int_{R^1} (|x|^p)^r \mu(dx)\right)^{1/r} \left(\int_{R^1} 1^s \mu(dx)\right)^{1/s}$$

$$= \left(\int_{R^1} |x|^q \mu(dx)\right)^{p/q}. \quad \blacksquare$$

$|M|_p(\mu)$ を与えると，分布 μ の散布の様子がある程度わかる．

定理 2.13
$$\mu([-a, a]) \geqq 1 - \frac{|M|_p(\mu)}{a^p}.$$

証明 $|M|_p(\mu) = \int_{R^1} |x|^p \mu(dx) \geqq \int_{[-a,a]^c} |x|^p \mu(dx) \geqq a^p \mu([-a, a]^c)$,

$\mu(\boldsymbol{R}^1)=1$ であるから，上の不等式から，定理の不等式が直ちにでる．$\blacksquare$

$M_1(\mu)$ を μ の**平均値**といい，$M(\mu)$ であらわし，

$$V(\mu) = \int_{R^1} (x - M(\mu))^2 \mu(dx)$$

を μ の**分散**という．Gauss 分布 $N_{m,v}$ の平均値は m，分散は v である．

d) たたみこみ

μ_1, μ_2 を分布とするとき

$$\mu(E) = \int_{R^1} \mu_1(E-x) \mu_2(dx), \qquad E \in \mathcal{B}^1$$

も分布となる．ここで

$$E - x = \{y - x \mid y \in E\}.$$

この分布 μ を μ_1, μ_2 の**たたみこみ** (convolution) といい，$\mu_1 * \mu_2$ であらわす．この定義が意味をもつためには，まず上の式の右辺の積分が意味をもつこと，次に μ が分布の条件を満足することを証明しなければならない．μ_1 が分布であるから，集合族

$$\mathcal{A} = \{E \in \mathcal{B}^1 \mid \mu_1(E-x) \text{ が } x \text{ に関して Borel 可測}\}$$

は Dynkin 族である．また

$$\mu_1((a, b] - x) = \mu_1(a-x, b-x]$$

は x に関して左連続であるから，Borel 可測である．したがって Dynkin 族定理により，$\mathcal{A}$ は $\mathcal{B}^1$ と一致する．これはすべての $E \in \mathcal{B}^1$ に対して $\mu_1(E-x)$ が x に関して Borel 可測であることを意味する．故に上の積分は意味をもち，$\mu(E)$, $E \in \mathcal{B}^1$, は定義可能である．これが分布の条件を満足することは容易にわかる．

直積測度に関する Fubini の定理により，$\mu_1 * \mu_2$ はまたつぎのようにかける．

$$(\mu_1 * \mu_2)(E) = \int_{R^2} 1_E(x+y) \mu_1(dx) \mu_2(dy).$$

$f: \boldsymbol{R}^1 \to \boldsymbol{R}^1$ を Borel 可測とすれば，$f(x+y)$ は $(x, y) \in \boldsymbol{R}^2$ の Borel 可測関数である．なぜならば $(x, y) \to x+y$ は Borel 可測関数 $g: \boldsymbol{R}^2 \to \boldsymbol{R}^1$ を定義し，$(x, y) \to f(x+y)$ は $f \circ g$ に等しいからである．任意の Borel 可測関数は，1_E, $E \in \mathcal{B}^1$, の1次結合の極限であらわされるから，すべての有界 Borel 関数 f に対し，上の式から，

$$\int_{R^1} f(x) \mu(dx) = \iint_{R^2} f(x+y) \mu_1(dx) \mu_2(dy)$$

が得られる．逆にこの式をみたす μ は $\mu_1 * \mu_2$ に等しいことは，$f = 1_E$ $(E \in \mathcal{B}^1)$ とおいてみれば，すぐにわかる．このことから

$$\mu_1 * \mu_2 = \mu_2 * \mu_1, \qquad (\mu_1 * \mu_2) * \mu_3 = \mu_1 * (\mu_2 * \mu_3)$$

が証明される．この式を示すには Fubini の定理を利用する必要があることに注意せよ．

上のようにたたみこみは交換律，結合律をみたすから，括弧をはずして

$$\mu_1 * \mu_2 * \cdots * \mu_n$$

とかいても混乱はおこらない．n に関する帰納法により，$\mu = \mu_1 * \mu_2 * \cdots * \mu_n$ は

$$\int_{R^1} f(x) \mu(dx) = \iint \cdots \int_{R^n} f(x_1 + x_2 + \cdots + x_n) \mu_1(dx_1) \mu_2(dx_2) \cdots \mu_n(dx_n)$$

(f は有界 Borel 関数)

と同等であることがわかる．

例題 2.5 (i) 次の関数はいずれもある分布の密度となっていることを示し，かつそれぞれの分布の分布関数のグラフをかけ ($-\infty < a < b < \infty$, $c > 0$, $\lambda > 0$).

$$U_{a,b}(x) = \begin{cases} (b-a)^{-1} & (a < x < b), \\ 0 & (\text{外では}) \end{cases} \qquad \text{(一様分布)},$$

$$T_{a,c}(x) = \begin{cases} \dfrac{c - |a-x|}{c^2} & (a-c < x < a+c), \\ 0 & (\text{外では}) \end{cases} \qquad \text{(三角分布)},$$

$$E_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geqq 0), \\ 0 & (\text{外では}) \end{cases} \qquad \text{(指数分布)}.$$

(ii) 定理 2.10 (i)$\Leftrightarrow$(iv) を利用して，次の極限関係を証明せよ．

$$\begin{pmatrix} k/n \\ 1/n \end{pmatrix}_{k=1,2,\cdots,n} \longrightarrow U_{0,1} \qquad (n\to\infty),$$

$$\begin{pmatrix} k/n \\ (\lambda/n)(1-\lambda/n)^{n-k} \end{pmatrix}_{k=1,2,\cdots} \longrightarrow E_\lambda \qquad (n\to\infty),$$

$$U_{-1/n,1/n} \longrightarrow \delta \qquad (n\to\infty).$$

(iii)　$\sup_{\mu\in\mathcal{M}}|M|_p(\mu)<\infty$ のときには，$\mathcal{M}$ の任意の無限列は収束部分列をもつことを示せ．（定理 2.13 と定理 2.11 を用いよ．）

(iv)　μ_1 が密度 f_1 をもてば，$\mu_1*\mu_2$ は密度

$$f(x) = \int_{R^1} f_1(x-y)\mu_2(dy)$$

をもち，また μ_1, μ_2 がそれぞれ密度 f_1, f_2 をもてば，$\mu_1*\mu_2$ は密度

$$f(x) = \int_{R^1} f_1(x-y)f_2(y)dy$$

をもつことを証明せよ．

(v)　$U_{-c,c}*U_{a,b}$ の密度のグラフをかけ．

(vi)　Gauss 分布，Cauchy 分布（例題 2.1(i)）について

$$N_{m,1/n} \longrightarrow \delta_m, \qquad C_{m,1/n} \longrightarrow \delta_m$$

を証明せよ．（この意味で $N_{m,0}=\delta_m$，$C_{m,0}=\delta_m$ と定義する．）

(vii)　$N_{m_1,v_1}*N_{m_2,v_2}=N_{m_1+m_2,v_1+v_2}$ を証明せよ．

(viii)　$C_{m_1,c_1}*C_{m_2,c_2}=C_{m_1+m_2,c_1+c_2}$ を証明せよ．

［ヒント］　複素関数論の留数定理を用いて定積分

$$\int_{-\infty}^{\infty} \frac{dy}{((x-y-m_1)^2+c_1{}^2)((y-m_2)^2+c_2{}^2)}$$

を計算せよ．（積分路 $-a\to a\to a+ia\to -a+ia\to -a$ に沿っての積分を考え，$a\to\infty$ のときの極限をとれ．）

(ix)　$r_n=\sum_{k=0}^{n} p_k q_{n-k}$ ならば

$$\begin{pmatrix} n \\ p_n \end{pmatrix}_{n=0,1,\cdots} * \begin{pmatrix} n \\ q_n \end{pmatrix}_{n=0,1,\cdots} = \begin{pmatrix} n \\ r_n \end{pmatrix}_{n=0,1,\cdots}$$

を証明し，これを用いて，2項分布，Poisson 分布について

$$b_{n,p}*b_{m,p} = b_{n+m,p}, \qquad p_\lambda * p_\mu = p_{\lambda+\mu}$$

を証明せよ．

§2.6　特性関数

a)　定義と例

μ を 1 次元の分布としたとき，関数

$$\varphi(z) = \int_{R^1} e^{izx}\mu(dx)$$

を μ の**特性関数** (characteristic function) または μ の **Fourier 変換**といい，$\mathcal{F}\mu$ であらわす．ある分布の特性関数を単に**特性関数**という．

$$\mu = \begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$$

のときには，

$$\mathcal{F}\mu(z) = \sum_n p_n e^{ia_n z},$$

したがって Poisson 分布 p_λ の特性関数は

$$\sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}e^{izn} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\lambda e^{iz})^n}{n!} = e^{-\lambda}e^{\lambda e^{iz}} = \exp\{\lambda(e^{iz}-1)\}.$$

また μ が密度 f をもつときには，

$$\mathcal{F}\mu(z) = \int_{R^1} e^{izx}f(x)dx.$$

したがって一様分布 $U_{a,b}$ の特性関数は

$$\frac{1}{b-a}\int_a^b e^{izx}dx = \frac{e^{izb}-e^{iza}}{iz(b-a)} = e^{iz(a+b)/2}\frac{e^{iz(b-a)/2}-e^{-iz(b-a)/2}}{2iz(b-a)/2}$$

$$= e^{iz(a+b)/2}\frac{\sin z(b-a)/2}{z(b-a)/2}.$$

また Gauss 分布 $N_{m,v}$ の特性関数は

$$\int_{-\infty}^{\infty} e^{izx}N_{m,v}(dx) = \exp\left\{imz-\frac{v}{2}z^2\right\}$$

である．これを証明しよう．$v=0$ のときには，$N_{m,0}=\delta_m$ であるから（例題 2.5(vi)），これは明らかである．$v>0$ のときには，上の積分で，$x=m+\sqrt{v}\,y$ とおいて

$$e^{izm}\int_{-\infty}^{\infty} e^{i(z\sqrt{v})y}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy$$

を得るから，

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{izy}e^{-y^2/2}dy = e^{-z^2/2}$$

をいえばよい．左辺は

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}e^{-(y+iz)^2/2}dy = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}\lim_{a\to\infty}\int_{-a+iz}^{a+iz} e^{-y^2/2}dy,$$

複素関数論の Cauchy の定理を用いると，これは

$$\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\lim_{a\to\infty}\int_{-a}^{a} e^{-y^2/2}dy = e^{-z^2/2}$$

に等しい．(積分路 $-a\to a\to a+iz\to -a+iz\to -a$ に沿っての積分を考えよ.)

μ が分布ならば，

$$\check{\mu}(E) = \mu(-E), \qquad -E = \{-x \mid x\in E\}$$

も分布である．これを μ の**裏がえし**という．このとき，有界な Borel 関数 f に対して

$$\int_{-\infty}^{\infty} f(x)\,\check{\mu}(dx) = \int_{-\infty}^{\infty} f(-x)\,\mu(dx)$$

がなりたつことは常套手段で容易に証明できる．特に $f(x)=e^{izx}$ とおいて

$$\mathcal{F}\check{\mu}(z) = \mathcal{F}\mu(-z) = \overline{\mathcal{F}\mu(z)}$$

が得られる．

さらに一般に $L_{ab}(x)=ax+b$ $(a\neq 0)$ による分布 μ の像測度 $\mu L_{ab}{}^{-1}$ もまた分布で

$$\int_{R^1} f(x)(\mu L_{ab}{}^{-1})(dx) = \int_{R^1} f(ax+b)\,\mu(dx),$$

したがって

$$\mathcal{F}(\mu L_{ab}{}^{-1})(z) = e^{ibz}(\mathcal{F}\mu)(az)$$

となることがわかる．$\check{\mu}$ は $\mu L_{ab}{}^{-1}$ で $a=-1$, $b=0$ の特別の場合である．

$\mu=\mu_1*\mu_2$ のときには

$$\int_{R^1} f(x)\,\mu(dx) = \int\int_{R^2} f(x+y)\,\mu_1(dx)\,\mu_2(dy),$$

$f(x)=e^{izx}$ とおいて

$$\mathcal{F}(\mu_1*\mu_2) = \mathcal{F}\mu_1\mathcal{F}\mu_2$$

が得られる．また分布の凸結合 $c_1\mu_1+c_2\mu_2$ $(c_1, c_2\geqq 0,\ c_1+c_2=1)$ に対しては

$$\mathcal{F}(c_1\mu_1+c_2\mu_2) = c_1\mathcal{F}\mu_1+c_2\mathcal{F}\mu_2.$$

$\mu_n\to\mu$ のときには，有界連続実関数 f に対して

$$\int_{R^1} f(x)\,\mu_n(dx) \longrightarrow \int_{R^1} f(x)\,\mu(dx)$$

となるが，これは f が複素数値でもよいことは，f を実部と虚部に分けて考えれば，明らかである．$f(x)=e^{izx}$ とおいて

$$\mu_n \longrightarrow \mu \implies \mathcal{F}\mu_n(z) \longrightarrow \mathcal{F}\mu(z)$$

となる．

さて $\varphi(z)=\mathcal{F}\mu(z)$ を与えると，逆に μ が定まる．すなわち

$$\mu \longrightarrow \mathcal{F}\mu$$

なる対応は1対1である．実際次の定理は μ を $\varphi=\mathcal{F}\mu$ から定める式を与えている．

定理 2.14 (Lévy の反転定理) $\varphi=\mathcal{F}\mu$ ならば

$$\mu(a,b)+\frac{1}{2}[\mu\{a\}+\mu\{b\}] = \lim_{c\to\infty}\frac{1}{2\pi}\int_{-c}^{c}\frac{e^{-ibz}-e^{-iaz}}{-iz}\varphi(z)\,dz.$$

証明 上の積分を $F(c)$ とおき，Fubini の定理により積分順序を交換して

$$\begin{aligned}
F(c) &= \int_{-c}^{c}\int_{a}^{b} e^{-ixz}dx\int_{-\infty}^{\infty} e^{iyz}\mu(dy)\,dz \\
&= \int_{-\infty}^{\infty}\mu(dy)\int_{a}^{b}dx\int_{-c}^{c} e^{iz(y-x)}dz \\
&= 2\int_{-\infty}^{\infty}\mu(dy)\int_{a}^{b}\frac{\sin c(x-y)}{x-y}dx \\
&= 2\int_{-\infty}^{\infty}\mu(dy)\int_{c(a-y)}^{c(b-y)}\frac{\sin u}{u}du \qquad \left(x=y+\frac{u}{c}\right) \\
&= 2\int_{-\infty}^{\infty}[G(c(b-y))-G(c(a-y))]\,\mu(dy),
\end{aligned}$$

$$G(x) = \int_{0}^{x}\frac{\sin u}{u}du.$$

$G(x)$ は x の連続関数で

$$\lim_{x\to\infty} G(x) = \frac{\pi}{2}, \qquad \lim_{x\to-\infty} G(x) = -\frac{\pi}{2}$$

であるから，$G(x)$ は $-\infty<x<\infty$ で有界である．したがって極限と積分とを交

換して

$$\lim_{c\to\infty} F(c) = 2\int_{-\infty}^{\infty} f_{a,b}(x)\mu(dy),$$

ここに

$$f_{a,b}(x) = \begin{cases} \dfrac{\pi}{2} & (x=a,b), \\ \pi & (a<x<b), \\ 0 & (\text{外では}), \end{cases}$$

したがって

$$\lim_{c\to\infty} F(c) = 2\pi\mu(a,b) + \pi\mu\{a\} + \pi\mu\{b\}. \qquad \blacksquare$$

上の定理により，a, b が μ の連続点ならば，$\mu(a,b)$ は $\varphi=\mathscr{F}\mu$ で定まり，したがって μ が φ で完全に定まる．故に対応 $\mu\to\mathscr{F}\mu$ は1対1である．これから分布をその特性関数で定義することもできる．例えば Gauss 分布 $N_{m,v}$ は $\exp\{imz-vz^2/2\}$ を特性関数とする分布であるといってもよい．同様に μ の裏がえし $\check{\mu}$ は $\overline{\mathscr{F}\mu}$ を特性関数とする分布である．分布と特性関数との関係を次頁の表に示しておこう．

$\mu_1 * \mu_2=\mu$ と $\mathscr{F}\mu_1\mathscr{F}\mu_2=\mathscr{F}\mu$ とは同等であるから，表から直ちに

$$b_{n,p} * b_{m,p} = b_{n+m,p},$$
$$p_\lambda * p_\mu = p_{\lambda+\mu},$$
$$\delta_{m_1} * \delta_{m_2} = \delta_{m_1+m_2},$$
$$U_{-c/2,c/2} * U_{-c/2,c/2} = T_{0,c},$$
$$N_{m_1,v_1} * N_{m_2,v_2} = N_{m_1+m_2,v_1+v_2},$$
$$C_{m_1,c_1} * C_{m_2,c_2} = C_{m_1+m_2,c_1+c_2}$$

が得られる．

b) 特性関数の性質

φ を分布 μ の特性関数とする，すなわち

$$\varphi(z) = \int_{-\infty}^{\infty} e^{izx}\mu(dx).$$

この定義から直ちに

$$|\varphi(z)| \leqq 1 = \varphi(0), \qquad \varphi(-z) = \overline{\varphi(z)}$$

分　　布	特性関数	
$\check{\mu}$	$\bar{\varphi}$	$(\varphi=\mathscr{F}\mu)$
$\mu*\check{\mu}$	$\lvert\varphi\rvert^2$	(〃)
$\mu L_{ab}{}^{-1}$ $(a\neq b)$	$e^{ibz}\varphi(az)$	(〃)
$c_1\mu_1+c_2\mu_2$ (凸結合)	$c_1\varphi_1+c_2\varphi_2$	$(\varphi_i=\mathscr{F}\mu_i)$
$\mu_1*\mu_2$	$\varphi_1\varphi_2$	(〃)
$\begin{pmatrix} a_n \\ p_n \end{pmatrix}_{n=1,2,\cdots}$	$\sum_n p_n e^{ia_nz}$	
$\mu(dx)=f(x)\,dx$	$\int_{-\infty}^{\infty} e^{ixz}f(x)\,dx$	
$\mu(dx)=f(ax+b)\lvert a\rvert dx$ $(a\neq 0)$	$e^{-izb/a}\varphi(z/a)$	$(\varphi=\mathscr{F}(f\cdot dx))$
$b_{n,p}$ (2項分布)	$(pe^{iz}+(1-p))^n$	
p_λ (Poisson 分布)	$\exp\{\lambda(e^{iz}-1)\}$	
δ_m (δ 分布)	e^{imz}	
$U_{a,b}$ (一様分布)	$(e^{ibz}-e^{iaz})/iz(b-a)$ $= e^{i(a+b)z/2}\dfrac{\sin(b-a)z/2}{(b-a)z/2}$	
$T_{a,c}$ (三角分布)	$e^{iaz}\left(\dfrac{\sin zc/2}{zc/2}\right)^2$	
$N_{m,v}$ (Gauss 分布)	$\exp\left\{imz-\dfrac{1}{2}vz^2\right\}$	
$C_{m,c}$ (Cauchy 分布)	$\exp\{imz-c\lvert z\rvert\}$	
E_λ (指数分布)	$(1-iz/\lambda)^{-1}$	

がでる．$|e^{izx}|\leqq 1$ であるから，有界収束定理により，

$$z_n \longrightarrow z \implies \varphi(z_n) \longrightarrow \varphi(z).$$

すなわち φ は連続である．さらに進んで φ が一様連続であることが次の不等式から得られる．

定理 2.15　$$|\varphi(z+h)-\varphi(z)| \leqq \sqrt{2|\varphi(0)-\varphi(h)|}.$$

証明　Schwarz の不等式を用いて

$$|\varphi(z+h)-\varphi(z)|^2 \leqq \int_{-\infty}^{\infty} |e^{i(z+h)x}-e^{izx}|^2\mu(dx)$$

$$= \int_{-\infty}^{\infty} 2(1-\cos hx)\,\mu(dx)$$

$$= 2\,\mathrm{Re}(\varphi(0)-\varphi(h)) \leqq 2|\varphi(0)-\varphi(h)|. \qquad \blacksquare$$

特性関数は正定 (positive definite) である，すなわち

$$\sum_{j,k=1}^{n} \xi_j\bar{\xi}_k\varphi(z_j-z_k) = \int_{-\infty}^{\infty}\left|\sum_{j=1}^{n}\xi_j e^{iz_jx}\right|^2\mu(dx) \geqq 0,$$

$$n=1,2,\cdots,\qquad \xi_1,\xi_2,\cdots,\xi_n\in \boldsymbol{C}^1,\qquad z_1,z_2,\cdots,z_n\in\boldsymbol{R}^1.$$

$\varphi(-z)=\overline{\varphi(z)}$ であるから，$(\varphi(z_j-z_k))_{j,k=1}^{n}$ は Hermite 行列であるが，上の不等式はこの行列が正定であることを示している．

c) 分布の収束と特性関数

分布 μ_n の特性関数を φ_n とする．$\mu_1,\mu_2,\cdots$ の収束と $\varphi_1,\varphi_2,\cdots$ の収束との関係を調べよう．

$$\mu_n \longrightarrow \mu$$

とすれば，定義により任意の有界連続複素関数 $f(x)$ に対して

$$\int_{\boldsymbol{R}^1} f(x)\,\mu_n(dx) \longrightarrow \int_{\boldsymbol{R}^1} f(x)\,\mu(dx).$$

特に $f(x)=e^{izx}$ とおいて，各点 $z\in\boldsymbol{R}^1$ に対し

$$\varphi_n(z)\longrightarrow\varphi(z)\equiv\mathscr{F}\mu(z).$$

実はこの収束は z に関して広義一様であることが証明できる (定理 2.16)．これをいうためにまず補題から始める．

補題 2.5(一般化された部分積分の定理) 任意の μ と任意の連続的微分可能な関数 f に対し

$$\int_{(a,b]} f(x)\,\mu(dx) = f(b)F(b)-f(a)F(a)-\int_{(a,b]} f'(x)F(x)\,dx.$$

ここで F は μ の分布関数である．

証明 $x,y\in(a,b]$ に対し

$$1_{(a,x]}(y) = 1_{[y,b]}(x) \qquad (a<x\leqq y\leqq b)$$

したがって $\boldsymbol{R}^1$ の上の任意の有界正則測度 μ,ν と $\boldsymbol{R}^2$ 上の有界 Borel 関数 $f(x,y)$ に対し

$$\int_{(a,b]}\int_{(a,x]} f(x,y)\,\mu(dy)\,\nu(dx)$$

$$= \int_{(a,b]}\int_{(a,b]} f(x,y)\,1_{(a,x]}(y)\,\mu(dy)\,\nu(dx).$$

上の等式と Fubini の定理により

$$= \int_{(a,b]}\int_{(a,b]} f(x,y)\,1_{[y,b]}(x)\,\nu(dx)\,\mu(dy)$$

$$= \int_{(a,b]}\int_{[y,b]} f(x,y)\,\nu(dx)\,\mu(dy).$$

この積分順序交換の式を頭において，定理の等式の左辺を変形する．

$$\int_{(a,b]} f(x)\,\mu(dx) = \int_{(a,b]}\int_{(a,x]} f'(y)\,dy\,\mu(dx)+\int_{(a,b]} f(a)\,\mu(dx)$$

$$= \int_{(a,b]}\int_{[y,b]}\mu(dx)f'(y)\,dy+f(a)\int_{(a,b]}\mu(dx)$$

$$= \int_{(a,b]}(F(b)-F(y-))f'(y)\,dy+f(a)(F(b)-F(a))$$

$$= f(b)F(b)-f(a)F(a)-\int_{(a,b]}F(y-)f'(y)\,dy.$$

$F(y-)$ は可算個の点 y を除いて $F(y)$ に等しいから，最後の式は定理の等式の右辺と一致する． $\blacksquare$

定理 2.16 $\varphi_n=\mathscr{F}\mu_n\ (n=1,2,\cdots)$，$\varphi=\mathscr{F}\mu$ としたとき，$\mu_n\to\mu$ ならば，z に関して広義一様に

$$\varphi_n(z)\longrightarrow\varphi(z).$$

証明 μ_n,μ の分布関数をそれぞれ F_n,F とすると，$\mu_n\to\mu$ から，F の連続点 x に対し

$$F_n(x)\longrightarrow F(x).$$

任意の $\varepsilon>0$ に対し，a を十分大きくとると

$$F(a)-F(-a)>1-\varepsilon.$$

この $-a,a$ がともに F の連続点になるように，a をとることもできる．そのようにすれば，十分大きい n に対し $(n>N_1=N_1(\varepsilon))$

$$F_n(a)-F_n(-a)>1-\varepsilon.$$

したがって $n>N_1(\varepsilon)$ に対し

$$\left|\varphi_n(z)-\int_{(-a,a]}e^{izx}\mu_n(dx)\right|\leqq\mu_n((-a,a]^c)=1-\mu_n(-a,a]<\varepsilon,$$

同様に

$$\left|\varphi(z)-\int_{(-a,a]} e^{izx}\mu(dx)\right|<\varepsilon.$$

上の補題により

$$\begin{aligned}
I_n(z) &= \int_{(-a,a]} e^{izx}\mu_n(dx) \\
&= e^{iza}F_n(a)-e^{-iza}F_n(-a)-\int_{(-a,a]} ize^{izx}F_n(x)\,dx, \\
I(z) &= \int_{(-a,a]} e^{izx}\mu(dx) \\
&= e^{iza}F(a)-e^{-iza}F(-a)-\int_{(-a,a]} ize^{izx}F(x)\,dx, \\
|I_n(z)-I(z)| &\leqq |F_n(a)-F(a)|+|F_n(-a)-F(-a)| \\
&\quad +\int_{(-a,a]} |z||F_n(x)-F(x)|\,dx.
\end{aligned}$$

右辺の第1項，第2項が $n\to\infty$ のとき0に近づくことは，$a, -a$ が F の連続点であることから明らか．また F の不連続点の集合は可算，したがって Lebesgue 測度0であるから，第3項の積分が $n\to\infty$ のとき0に近づくことは，有界収束定理からでる．しかもその収束は $|z|\leqq b$ (b は任意の定数) で一様であるから

$$\limsup_{n\to\infty}\sup_{|z|\leqq b}|\varphi_n(z)-\varphi(z)|\leqq 2\varepsilon.$$

ε は任意の正数であるから，定理は完全に証明された．▌

上の定理で分布の収束から対応する特性関数の収束を導いたが，逆に特性関数の収束から分布の収束を導く次の定理を証明しよう．

定理 2.17 (**Glivenko の定理**) φ_n, φ をそれぞれ分布 μ_n, μ の特性関数とする．

$$\varphi_n(z)\longrightarrow\varphi(z)\ \text{(各点収束)} \Longrightarrow \mu_n\longrightarrow\mu.$$

証明 g を $\boldsymbol{R}^1$ の上で可積分とする．(以下積分はすべて $\boldsymbol{R}^1$ 上の積分とする.)

$$\begin{aligned}
\int\!\!\int e^{izx}g(z)\,dz\mu_n(dx) &= \int\!\!\int e^{izx}\mu_n(dx)g(z)\,dz \qquad \text{(Fubini の定理)} \\
&= \int \varphi_n(z)g(z)\,dz,
\end{aligned}$$

$n\to\infty$ として

$$\begin{aligned}
&\longrightarrow \int\varphi(z)g(z)\,dz \qquad \text{(有界収束定理)} \\
&= \int\!\!\int e^{izx}\mu(dx)g(z)\,dz \\
&= \int\!\!\int e^{izx}g(z)\,dz\mu(dx) \qquad \text{(Fubini の定理)}.
\end{aligned}$$

したがって

$$h(x)=\int e^{izx}g(z)\,dz \qquad (g\ \text{は}\ \boldsymbol{R}^1\ \text{で可積分})$$

に対して

$$\int h(x)\mu_n(dx)\longrightarrow\int h(x)\mu(dx).$$

したがってコンパクトな台をもつ連続関数 f がこのような h で一様近似できれば，f に対しても

$$\int f(x)\mu_n(dx)\longrightarrow\int f(x)\mu(dx)$$

となり，定理 2.10 (i)$\Longleftrightarrow$(ii) により証明は完了する．

f に対して

$$g_n(z)=\frac{1}{2\pi}e^{-z^2/2n}\int e^{-izy}f(y)\,dy$$

とおくと，

$$|g_n(z)|\leqq\frac{1}{2\pi}e^{-z^2/2n}\int|f(x)|\,dx$$

であるから，g_n は $\boldsymbol{R}^1$ 上で可積分である．さて

$$h_n(x)=\int e^{izx}g_n(z)\,dz\xrightarrow[(n\to\infty)]{} f(x) \qquad \text{(一様収束)}$$

を証明しよう．

$$\begin{aligned}
h_n(x) &= \int e^{izx}\frac{1}{2\pi}e^{-z^2/2n}\int e^{-izy}f(y)\,dy\,dz \\
&= \frac{1}{2\pi}\int\!\!\int e^{iz(x-y)}e^{-z^2/2n}dz f(y)\,dy \qquad \text{(Fubini の定理)}
\end{aligned}$$

$$= \sqrt{\frac{n}{2\pi}} \int\int e^{iz(x-y)} \frac{1}{\sqrt{2\pi n}} e^{-z^2/2n} dz f(y) dy$$

$$= \sqrt{\frac{n}{2\pi}} \int e^{-n(x-y)^2/2} f(y) dy \qquad (\mathscr{F}N_{0,n}(z) = e^{-nz^2/2})$$

$$= \frac{1}{\sqrt{2\pi}} \int e^{-t^2/2} f\left(x + \frac{t}{\sqrt{n}}\right) dt,$$

f は明らかに連続かつ有界であるから，上の式で $n\to\infty$ とすると，有界収束定理により，$h_n(x)\to f(x)$ が各点 x でなりたつことがわかる．この収束の一様性をいうには，もう少し詳しい評価をしなければならない．

まず明らかに

$$f(x) = \frac{1}{\sqrt{2\pi}} \int e^{-t^2/2} f(x) dt$$

であるから

$$|h_n(x) - f(x)| \leqq \frac{1}{\sqrt{2\pi}} \int e^{-t^2/2} \left| f\left(x + \frac{t}{\sqrt{n}}\right) - f(x) \right| dt$$

$$\leqq \frac{1}{\sqrt{2\pi}} \left[\sup_{|t|\leqq a} \left| f\left(x + \frac{t}{\sqrt{n}}\right) - f(x) \right| \int_{|t|\leqq a} e^{-t^2/2} dt \right.$$

$$\left. + 2b \int_{|t|>a} e^{-t^2/2} dt \right] \qquad (b = \sup_x |f(x)| < \infty),$$

$$\sup_x |h_n(x) - f(x)| \leqq \sup_{\substack{|t|\leqq a \\ x\in R^1}} \left| f\left(x + \frac{t}{\sqrt{n}}\right) - f(x) \right| + 2b \int_{|t|>a} e^{-t^2/2} dt,$$

$f(x)$ は一様連続であるから，右辺の第1項は $n\to\infty$ のとき0となる．

$$\limsup_{n\to\infty} [\sup_x |h_n(x) - f(x)|] \leqq 2b \int_{|x|>a} e^{-t^2/2} dt \longrightarrow 0 \qquad (a\to\infty). \quad \blacksquare$$

φ_n を分布 μ_n の特性関数とし，各点で $\varphi_n(z)$ がある関数 $\varphi(z)$ に収束するとする．もし極限関数 φ がある分布 μ の特性関数であれば，上の Glivenko の定理により μ_n が μ に収束することがわかるが，一般には φ は特性関数ではない．これは次の例を見ればすぐにわかる．

$\mu_n = N_{0,n}$ とすれば，

$$\varphi_n(z) = \mathscr{F}N_{0,n}(z) = e^{-nz^2/2} \longrightarrow \varphi(z) = \begin{cases} 1, & z = 0, \\ 0, & z \neq 0 \end{cases}$$

であるが，φ は，$z=0$ で連続ではないから，特性関数ではない．しかし次の定理がなりたつ．

定理 2.18 (Lévy の収束定理) $\varphi_n = \mathscr{F}\mu_n$ が各点で φ に収束し，しかもこの収束が0のある近傍 $|z|<a$ で (a はどれ程小さくてもよい) 一様であれば，極限関数 φ はある分布 μ の特性関数である．したがって Glivenko の定理により $\mu_n\to\mu$ を得る．

証明 φ_n は連続であるから，一様収束の仮定により，$\varphi(z)$ は $z=0$ で連続である．明らかに

$$\varphi(0) = \lim_{n\to\infty} \varphi_n(0) = 1$$

であるから，正数 $\delta=\delta(\varepsilon)<a$ を適当にとって，$|z|<\delta$ なる限り

$$|\varphi(z) - 1| < \varepsilon$$

とできる．$\varphi_n(z)$ は $|z|<\delta$ で $\varphi(z)$ に一様収束するから，$N=N(\varepsilon)$ (z には無関係) を十分大きくとると，$n>N$，$|z|<\delta$ なる限り

$$|\varphi_n(z) - 1| < \varepsilon.$$

したがって

$$1-\varepsilon < \frac{1}{2\delta} \int_{-\delta}^{\delta} \mathrm{Re}(\varphi_n(z)) dz = \frac{1}{2\delta} \int_{-\delta}^{\delta} \int_{-\infty}^{\infty} (\cos zx) \mu_n(dx) dz$$

$$= \int_{-\infty}^{\infty} \frac{\sin \delta x}{\delta x} \mu_n(dx),$$

被積分関数を $[-b, b]$ では1でおさえ，外では $1/\delta b$ でおさえて

$$\leqq \mu_n[-b, b] + \frac{1}{\delta b} \qquad (b \text{ は任意の正数}).$$

故に

$$\mu_n[-b, b] \geqq 1 - \varepsilon - \frac{1}{\delta b},$$

$$\inf_{n>N} \mu_n[-b, b] \geqq 1 - \varepsilon - \frac{1}{\delta b},$$

$$\lim_{b\to\infty} \inf_{n>N} \mu_n[-b, b] \geqq 1 - \varepsilon.$$

さておのおのの n に対しては当然

$$\lim_{b\to\infty} \mu_n[-b, b] = 1$$

から

$$\lim_{b\to\infty}\inf_{n\le N}\mu_n[-b,b]=1\geqq 1-\varepsilon,$$

したがって

$$\lim_{b\to\infty}\inf_{n}\mu_n[-b,b]\geqq 1-\varepsilon\longrightarrow 1 \qquad (\varepsilon\to 0).$$

これから，定理2.11により $\{\mu_n\}$ は収束部分列 $\{\nu_n\}$ をもつ．$\{\nu_n\}$ の極限分布を μ とすると

$$\mathscr{F}\nu_n(z)\longrightarrow\mathscr{F}\mu(z)$$

が各点で(実は広義一様に)なりたつ．$\mathscr{F}\nu_n$ は $\varphi_n=\mathscr{F}\mu_n$ の部分列であり，また

$$\varphi_n(z)\longrightarrow\varphi(z)$$

であるから，$\varphi=\mathscr{F}\mu$ となる．▌

d) Bochner の定理

複素数値関数 $\varphi(z)$ がどういう性質をもてば，特性関数となるかを示すのが Bochner の定理である．

定理 2.19 (**Bochner の定理**) $\varphi(z)$ が特性関数となるためには次の三つの条件がなりたつことが必要十分である．

(B. 1) φ は正定である，

(B. 2) $\varphi(z)$ は $z=0$ で連続である，

(B. 3) $\varphi(0)=1$.

証明 必要なことは明らかであるから，十分性だけ証明すればよい．φ が上の三つの条件をみたすとせよ．まず正定性から

$$\sum_{j,k=1}^{n}\xi_j\bar{\xi}_k\varphi(z_j-z_k)\geqq 0,$$

$$n=1,2,\cdots,\qquad \xi_1,\xi_2,\cdots,\xi_n\in \boldsymbol{C}^1,\qquad z_1,z_2,\cdots,z_n\in\boldsymbol{R}^1.$$

$n=2$, $\xi_1=\xi_2=1$, $z_1=z$, $z_2=0$ とおいて

$$2\varphi(0)+\varphi(z)+\varphi(-z)\geqq 0.$$

したがって $\varphi(z)+\varphi(-z)$ は実である．また $n=2$, $\xi_1=i$, $\xi_2=1$, $z_1=z$, $z_2=0$ とおいて

$$2\varphi(0)+i\varphi(z)-i\varphi(-z)\geqq 0.$$

したがって $\varphi(z)-\varphi(-z)$ は純虚である．このことから

$$\varphi(-z)=\overline{\varphi(z)}$$

を得る．これを頭において，$n=2$, $\xi_1=\xi$, $\xi_2=\eta$, $z_1=z$, $z_2=0$ とすると

$$\xi\bar{\xi}+\xi\bar{\eta}\varphi(z)+\bar{\xi}\eta\overline{\varphi(z)}+\eta\bar{\eta}\geqq 0.$$

故に

$$\begin{bmatrix}1 & \overline{\varphi(z)}\\ \varphi(z) & 1\end{bmatrix}\ \text{は正定行列}\quad\text{すなわち}\quad \begin{vmatrix}1 & \overline{\varphi(z)}\\ \varphi(z) & 1\end{vmatrix}\geqq 0$$

となる．これから

$$|\varphi(z)|^2\leqq 1\quad\text{すなわち}\quad|\varphi(z)|\leqq 1$$

を得て，φ が有界関数であることがわかる．また

$$n=2,\quad \xi_1=\xi,\ \xi_2=-\xi,\ \xi_3=\eta,\quad z_1=0,\ z_2=h,\ z_3=z+h$$

とおき，$\varphi(0)=1$ を考慮して整頓すると

$$\xi\bar{\xi}(2-\varphi(h)-\varphi(-h))+\xi\bar{\eta}(\overline{\varphi(z+h)}-\overline{\varphi(z)})+\bar{\xi}\eta(\varphi(z+h)-\varphi(z))+\eta\bar{\eta}\geqq 0.$$

これから上と同様に

$$|\varphi(z+h)-\varphi(z)|^2\leqq 2-\varphi(h)-\varphi(-h)=2\,\mathrm{Re}\,(1-\varphi(h))\leqq 2|1-\varphi(h)|.$$

φ は $z=0$ で連続であるから，上の式により φ は一様連続となることがわかる．

かくして φ が有界一様連続関数であることがわかった．故に連続，有界で可積分な関数 g に対して

$$\int\!\!\int\varphi(t-s)g(t)\overline{g(s)}\,dtds$$
$$=\lim_{n\to\infty}\sum_{j,k=-n^2}^{n^2}\varphi\Big(\frac{j}{n}-\frac{k}{n}\Big)g\Big(\frac{j}{n}\Big)\overline{g\Big(\frac{k}{n}\Big)}\Big(\frac{1}{n}\Big)^2\geqq 0,$$

すなわち

$$\int\varphi(t)\int g(t+s)\overline{g(s)}\,dsdt\geqq 0.$$

特に

$$g(t)=N_{0,n/4}(t)e^{-ixt}$$

とおくと，

$$N_{0,v}(t)=N_{0,v}(-t),\qquad N_{0,\alpha}*N_{0,\beta}=N_{0,\alpha+\beta}$$

により

$$\int g(t+s)\overline{g(s)}\,ds=N_{0,n/2}(t)e^{-ixt}=\sqrt{\frac{1}{\pi n}}e^{-t^2/n}e^{-ixt},$$

故に

$$\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}e^{-ixt}dt \geqq 0.$$

この左辺に $(2\pi)^{-1}$ をかけたものを $f_n(x)$ とおく.

$$\mu_n(dx) = f_n(x)dx$$

が分布を定めることをいう. μ_n は明らかに $\mathcal{B}^1$ 上の測度であるから, $\mu_n(\boldsymbol{R}^1)=1$ を示せばよい.

$$\mu_n(-a, a) = \frac{1}{2\pi}\int_{-a}^{a}f_n(x)dx$$
$$= \frac{1}{2\pi}\int_{-a}^{a}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}e^{-ixt}dtdx,$$

Fubini の定理を用いて積分順序を交換して

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}\frac{2\sin at}{t}dt,$$

$a\to\infty$ のとき $\mu_n(-a, a)\uparrow\mu_n(\boldsymbol{R}^1)$ であるから,

$$\mu_n(\boldsymbol{R}^1) = \lim_{b\to\infty}\frac{1}{b}\int_0^b\mu_n(-a, a)da$$
$$= \lim_{b\to\infty}\frac{1}{b}\int_0^b\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}\frac{2\sin at}{t}dtda$$
$$= \lim_{b\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}\frac{2(1-\cos tb)}{t^2b}dt$$
$$= \lim_{b\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi\left(\frac{t}{b}\right)e^{-t^2/nb^2}\frac{2(1-\cos t)}{t^2}dt.$$

ここで,

$$\frac{1-\cos t}{t^2} \geqq 0, \qquad \int_{-\infty}^{\infty}\frac{1-\cos t}{t^2}dt = \pi$$

であるから, 有界収束定理により, 上の極限は $\varphi(0)$ すなわち1に等しい. これで μ_n が分布であることがわかった.

次に分布 μ_n の特性関数 $\varphi_n(z)$ が $\varphi(z)e^{-z^2/n}$ に等しいことを示そう.

$$\varphi_n(z) = \int_{-\infty}^{\infty}e^{izx}f_n(x)dx$$

$$= \lim_{a\to\infty}\int_{-a}^{a}e^{izx}\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}e^{itx}dtdx$$
$$= \lim_{a\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}\frac{2\sin a(t-z)}{t-z}dt.$$

前と同様の考えで

$$= \lim_{b\to\infty}\frac{1}{b}\int_0^b da\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}\frac{2\sin a(t-z)}{t-z}dt$$
$$= \lim_{b\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)e^{-t^2/n}\frac{2(1-\cos b(t-z))}{b(t-z)^2}dt$$
$$= \lim_{b\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi\left(z+\frac{s}{b}\right)e^{-(z+s/b)^2/n}\frac{2(1-\cos s)}{s^2}ds$$
$$= \varphi(z)e^{-z^2/n}.$$

したがってすべての z に対し

$$\varphi_n(z) \longrightarrow \varphi(z) \qquad (n\to\infty)$$

で, この収束は $|z|\leqq 1$ で一様である. 故に Lévy の収束定理により, φ はある分布の特性関数である. ▮

例題 2.6 (i) $f(x)$ が $\boldsymbol{R}^1$ 上の可積分複素数値関数とするとき, その Fourier 変換 $\varphi=\mathcal{F}f$ を

$$\varphi(t) = \int_{-\infty}^{\infty}e^{ixt}f(x)dx$$

で定義する. もし f が連続, φ が可積分ならば

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ixt}\varphi(t)dt$$

がなりたつことを示せ.

[ヒント] Bochner の定理の証明で用いた考えにより

$$\int_{-\infty}^{\infty}e^{-ixt}\varphi(t)dt = \lim_{a\to\infty}\int_{-a}^{a}e^{-ixt}\varphi(t)dt$$
$$= \lim_{b\to\infty}\frac{1}{b}\int_0^b da\int_{-a}^{a}e^{-ixt}\varphi(t)dt.$$

(ii) $f(x)=(2\pi)^{-1}(\sin(x/2)/(x/2))^2$ がある分布の密度であることを示し, その分布の特性関数は $T_{0,1}$ に等しいことを示せ.

$$T_{0,1}(z)=\begin{cases}1-|z|, & |z|\leqq 1\\ 0, & |z|>1\end{cases}=(1-|z|)^+.$$

［ヒント］ $\mathcal{F}T_{0,1}(x)=((\sin 2^{-1}x)/2^{-1}x)^2$ に注意し，上の (i) を用いよ．

(iii) 確率空間 (A,ν) の点 a に依存する関数 $F_a(\xi), \xi\in\Xi$ があり，おのおのの ξ に対して $F_a(\xi)$ が a について可積分とする．このとき

$$F(\xi)=\int_A F_a(\xi)\nu(da), \qquad \xi\in\Xi$$

を関数族 $\{F_a, a\in A\}$ の**積分的凸結合**という．分布(Borel 集合の関数と見る)の積分的凸結合は分布であり，特性関数の積分的凸結合は特性関数であることを示せ．

(iv) (**Pólya の定理**) $\varphi(z)$ が非負偶関数で $z>0$ では減少かつ凸であり，$z=0$ で連続かつ $\varphi(0)=1$ とする．このとき φ は特性関数となる．これを証明せよ．

［ヒント］ 一般に $\psi(z)$ が特性関数ならば，$\psi(z/a)$ $(a>0)$ も特性関数であるから

$$\varphi_a(z)=\left(1-\left|\frac{z}{a}\right|\right)^+, \qquad 0<a<\infty$$

は特性関数である．また

$$\varphi_\infty(z)\equiv 1$$

も $\mathcal{F}\delta$ に等しく特性関数である．問題の $\varphi(z)$ が関数族 $\varphi_a(z)$, $0<a\leqq\infty$, の積分的凸結合であることを証明すればよい．結合の重みをあらわす $(0,\infty]$ の上の確率測度 α を求めるには

$$\varphi(z)=\int_{(0,\infty]}\varphi_a(z)\nu(da)$$

とおくと，$z>0$ では

$$\begin{aligned}\varphi(z)&=\int_{[z,\infty)}\left(1-\frac{z}{a}\right)\nu(da)+\nu\{\infty\}\\&=\int_{[z,\infty)}\int_{[z,a)}d\xi\frac{1}{a}\nu(da)+\nu\{\infty\}\\&=\int_{[z,\infty)}\int_{(\xi,\infty)}\frac{1}{a}\nu(da)d\xi+\nu\{\infty\},\end{aligned}$$

故に

$$\nu\{\infty\}=\lim_{z\to\infty}\varphi(z).$$

$0<z<\infty$ では

$$\varphi^+(z)\,(\varphi\text{ の右微分係数})=-\int_{[z,\infty)}\frac{1}{a}\nu(da),$$

$$d\varphi^+(z)=\frac{1}{z}\nu(dz),$$

$$\nu(dz)=zd\varphi^+(z)\geqq 0 \qquad (\varphi\text{ は凸であるから}).$$

これを頭において

$$\int_{(0,\infty)}ad\varphi^+(a)+\nu\{\infty\}=1,$$

$$\int_{(0,\infty)}\varphi_a(z)ad\varphi^+(a)+\nu\{\infty\}\varphi_\infty(z)=\varphi(z)$$

を証明すればよい．前者は後者で $z=0$ とおいたものであるから，後者だけ示せばよいが，それには $\varphi^+(\infty)=0$, $\varphi_\infty(z)\equiv 1$ を用い，かつ

$$\begin{aligned}\int_{(0,\infty)}\varphi_a(z)ad\varphi^+(a)&=\int_{[z,\infty)}(a-z)d\varphi^+(a)\\&=\int_{[z,\infty)}\int_{[z,a)}d\xi d\varphi^+(a)\end{aligned}$$

を考慮し，積分順序を交換せよ．

(v) 複素関数論の留数定理を用いて Cauchy 分布の特性関数を求めよ．

(vi) $\varphi_n(z)$ が特性関数で $\varphi_n(z)\to\varphi(z)$ (各点収束) とする．$\varphi(z)$ が $z=0$ で連続ならば，φ も特性関数であることを，Bochner の定理を用いて証明せよ．(この事実は Lévy の収束定理を含む.)

(vii) 分布 μ の n 次絶対能率 $|M|_n(\mu)$ が有限ならば，μ の特性関数 $\varphi(z)$ は n 回まで微分可能で

$$\varphi^{(k)}(z)=\int_{R^1}(ix)^k e^{izx}\mu(dx), \qquad k=0,1,2,\cdots,n \qquad (\varphi^{(0)}\equiv\varphi)$$

となることを示せ．

［ヒント］ 仮定により $|M|_k(\mu)$ は $k=1,2,\cdots,n$ に対し有限．したがって上の式の右辺の積分は有限で，z について連続(有界収束定理)．上の等式は $k=0$ に対しては明らか．k についてなりたてば，

$$\frac{1}{h}(\varphi^{(k)}(z+h)-\varphi^{(k)}(z)) = \int_{-\infty}^{\infty}\frac{e^{i(z+h)x}-e^{izx}}{h}(ix)^k\mu(dx)$$
$$= \frac{1}{h}\int_{-\infty}^{\infty}\int_z^{z+h} e^{i\zeta x}d\zeta\,(ix)^{k+1}\mu(dx)$$
$$= \frac{1}{h}\int_z^{z+h} d\zeta\int_{-\infty}^{\infty} e^{i\zeta x}(ix)^{k+1}\mu(dx)$$

(Fubini の定理)

$$\longrightarrow \int_{-\infty}^{\infty} e^{izx}(ix)^{k+1}\mu(dx) \qquad (h\to 0).$$

(viii)　分布 μ の特性関数 $\varphi(z)$ が $z=0$ で $2n$ 回微分可能ならば，$|M|_{2n}(\mu)<\infty$ で，$\varphi^{(k)}(z)$ $(k=0,1,2,\cdots,2n)$ は前問の式で与えられる．とくに $z=0$ とおいて

$$M_k(\mu) = i^{-k}\varphi^{(k)}(0), \qquad k = 0,1,2,\cdots,2n.$$

[ヒント]　$k=1$ のときには

$$-\frac{1}{h^2}(\varphi(h)-2\varphi(0)+\varphi(-h)) = \int_{-\infty}^{\infty}\frac{2(1-\cos hx)}{h^2}\mu(dx)$$
$$\geqq \int_{-A}^{A}\frac{2(1-\cos hx)}{h^2}\mu(dx)$$
$$= \frac{4}{h^2}\int_{-A}^{A}\left(\sin\frac{hx}{2}\right)^2\mu(dx)$$
$$\geqq \frac{4}{h^2}\int_{-A}^{A}\left(\frac{2}{\pi}\right)^2\left(\frac{hx}{2}\right)^2\mu(dx)$$

($|h|<\pi/A$ なる限り)

$$= \left(\frac{2}{\pi}\right)^2\int_{-A}^{A}x^2\mu(dx),$$

$\varphi''(0)$ が存在するから，左辺は $h\downarrow 0$ のとき，$\varphi''(0)$ に近づく．したがって

$$-\varphi''(0) \geqq 4\int_{-A}^{A}x^2\mu(dx),$$

$A\to\infty$ として

$$-\varphi''(0) \geqq 4\int_{-\infty}^{\infty}x^2\mu(dx).$$

これは $|M|_2(\mu)<\infty$ を意味する．したがって前問により

$$\varphi^{(2)}(z) = \int_{-\infty}^{\infty}e^{izx}(ix)^2\mu(dx) \qquad (z\in \boldsymbol{R}^1).$$

これに同様の論法を施して $|M|_4(\mu)<\infty$ を得る．同様に帰納法で $|M|_{2k}(\mu)<\infty$ $(k=1,2,\cdots,n)$ がいえる．

(ix)　前問を利用して，Gauss 分布 $N_{0,1}$ の k 次の能率は，k が奇数ならば 0，偶数 $(k=2p)$ ならば $1\cdot 3\cdot 5\cdot\cdots\cdot(2p-1)$ となることを示せ．

[ヒント]　$N_{0,1}$ の特性関数は

$$e^{-z^2/2} = \sum_{p=0}^{\infty}\left(-\frac{1}{2}\right)^p\frac{z^{2p}}{p!}.$$

§2.7　分布族の中の位相

1次元の分布の全体を $\mathcal{P}^1$ であらわす．$\mathcal{P}^1$ の中の列の極限については§2.5 b) で定義したが，本節では $\mathcal{P}^1$ の中に Hausdorff 位相をいれ，その位相による収束が前の収束と一致するようにしたい．C_b を $\boldsymbol{R}^1$ 上の有界連続実関数の全体とし，C_b の元のうちでコンパクトな台をもつものの全体を C_K であらわす．$f\in C_b$，$\mu\in\mathcal{P}^1$ に対し

$$(f,\mu) = \int_{\boldsymbol{R}^1}f(x)\mu(dx)$$

とおく．$\mu\in\mathcal{P}^1$ の近傍系として

$$U_{f_1,f_2,\cdots,f_n,\varepsilon}(\mu) = \{\nu\in\mathcal{P}^1\,|\,|(f_k,\nu)-(f_k,\mu)|<\varepsilon, k=1,2,\cdots,n\},$$
$$f_1,f_2,\cdots,f_n\in C_b, \qquad \varepsilon>0, \qquad n=1,2,\cdots$$

とすると，これが $\mathcal{P}^1$ の中の Hausdorff 位相を定めることは容易にわかる．この位相を $\mathcal{P}^1$ の中の**弱位相**という．$\{\mu_n\}$ を $\mathcal{P}^1$ の中の列とするとき，§2.5 b) の意味での '$\mu_n\to\mu$' がこの弱位相の意味での '$\mu_n\to\mu$' と一致することは明らかである．

C_b のかわりに C_K を用いて上の弱位相と同じように $\mathcal{P}^1$ の中の Hausdorff 位相が定義できる．この位相は一見弱位相より弱い位相であるが，実は同じ位相であることが次のようにしてわかる．それには任意の $\varepsilon>0$，$f_1,f_2,\cdots,f_n\in C_b$ に対して，$\delta>0$，$g_0,g_1,\cdots,g_n\in C_K$ を適当にとり，

$$U_{g_0,g_1,\cdots,g_n,\delta}(\mu)\subset U_{f_1,f_2,\cdots,f_n,\varepsilon}(\mu)$$

をいえば十分である．$e_A\in C_K$ を

$$e_A(x)=\begin{cases}1, & |x|\leqq A,\\ 0, & |x|\geqq A+1,\\ x\text{ の1次式}, & x\in[-A-1,-A]\text{ または } x\in[A,A+1]\end{cases}$$

で定めると，任意の $\nu\in\mathcal{P}^1$ に対し，

$$\begin{aligned}|(e_Af_k,\nu)-(f_k,\nu)| &= |((1-e_A)f_k,\nu)|\\ &\leqq |(\|f_k\|_\infty(1-e_A),\nu)| \qquad (\|f_k\|_\infty=\sup_x|f_k(x)|)\\ &\leqq \|f_k\|_\infty(1-e_A,\nu)\\ &\leqq \|f_k\|_\infty(1-(e_A,\mu))+\|f_k\|_\infty|(e_A,\nu)-(e_A,\mu)|\\ &\leqq \|f_k\|_\infty\mu([-A,A]^c)+\|f_k\|_\infty|(e_A,\nu)-(e_A,\mu)|\\ &\leqq a(\mu([-A,A]^c)+|(e_A,\nu)-(e_A,\mu)|) \qquad (a=\max_k\|f_k\|_\infty).\end{aligned}$$

とくに $\nu=\mu$ のときには

$$|(e_Af_k,\mu)-(f_k,\mu)|\leqq a\mu([-A,A]^c).$$

故に

$$\begin{aligned}|(f_k,\nu)-(f_k,\mu)|\leqq |(e_Af_k,\nu)-(e_Af_k,\mu)|&+2a\mu([-A,A]^c)\\ &+a|(e_A,\nu)-(e_A,\mu)|.\end{aligned}$$

さて $\varepsilon>0$ に対し，A を十分大きくとって

$$2a\mu([-A,A]^c)<\frac{\varepsilon}{2}$$

とすると，

$$|(e_A,\nu)-(e_A,\mu)|<\frac{\varepsilon}{4a},\qquad |(e_Af_k,\nu)-(e_Af_k,\mu)|<\frac{\varepsilon}{4},\qquad k=1,2,\cdots,n$$

なる限り

$$|(f_k,\nu)-(f_k,\mu)|<\varepsilon,\qquad k=1,2,\cdots,n.$$

これは $g_0=e_A$，$g_k=e_Af_k$ $(k=1,2,\cdots,n)$，$\delta=\varepsilon/(4a+4)$ に対し

$$U_{g_0,g_1,\cdots,g_n,\delta}(\mu)\subset U_{f_1,f_2,\cdots,f_n,\varepsilon}(\mu)$$

がなりたつことを意味する．g_k はすべて C_K に属するから，これで両位相の一致することが証明された．

次に弱位相は距離によって定義されることを示そう．C_K の中で台が $[-n,n]$ に含まれるものの全体を $C_K^{(n)}$ とすると，$C_K^{(n)}$ の中の関数は $[-n,n]$ の外で 0，その中では一様連続であるから，部分的1次の関数で一様近似され，しかも近似関数としては分点もその点における値も有理数にとることができるから，$C_K^{(n)}$ の中には一様収束の意味で稠密な可算部分集合がある．C_K は $C_K^{(n)}$ $(n=1,2,\cdots)$ の和集合であるから，C_K についても同様のことがいえる．C_K の中で稠密な可算部分集合を $g_1,g_2,\cdots$ とする．さて μ,ν の距離を

$$\rho(\mu,\nu)=\sum_{n=1}^\infty 2^{-n}(|(g_n,\mu)-(g_n,\nu)|\wedge 1),\qquad a\wedge b=\min(a,b)$$

と定義すると，ρ で定まる位相（ρ 位相）は弱位相と一致することが容易に証明できる．まず N を十分大きくとって

$$\sum_{n=N+1}^\infty 2^{-n}<\varepsilon$$

とし，

$$\nu\in U_{g_1,g_2,\cdots,g_N,\varepsilon}(\mu)$$

とすれば

$$\rho(\mu,\nu)<2\varepsilon$$

となるから，弱位相は ρ 位相よりも強い．逆に任意の $f_1,f_2,\cdots,f_n\in C_K$ に対して N を十分大きくとり，$g_1,g_2,\cdots,g_N$ の中に任意の f_k $(k=1,2,\cdots,n)$ を ε 近似するもの $g_{p(k)}$ があるようにすると，$\rho(\mu,\nu)<\varepsilon$ なるかぎり $(0<\varepsilon<1/2)$，

$$|(f_k,\nu)-(f_k,\mu)|\leqq 2\varepsilon+|(g_{p(k)},\nu)-(g_{p(k)},\mu)|.$$

$|(f_k,\nu)|,\ |(f_k,\mu)|\leqq\|f_k\|_\infty$ であるから，

$$\begin{aligned}|(f_k,\nu)-(f_k,\mu)|&\leqq 2\varepsilon+(2a+1)(|(g_{p(k)},\nu)-(g_{p(k)},\mu)|\wedge 1)\\ &\leqq 2\varepsilon+(2a+1)2^N\rho(\mu,\nu),\end{aligned}$$

$$a=\max_k\|f_k\|_\infty,\qquad k=1,2,\cdots,n.$$

これから $\rho(\mu,\nu)$ が十分小さいならば，ν は $U_{f_1,f_2,\cdots,f_n,3\varepsilon}(\mu)$ に属することがわかる．故に ρ 位相は弱位相よりも強い．

$\mathcal{P}^1$ の中の弱位相は距離 ρ で定まることがわかった．これから弱位相は第1可算性をもつことがわかる．実は第2可算性をもつのである．それには

$$\begin{pmatrix}a_1, a_2, \cdots, a_n\\ p_1, p_2, \cdots, p_n\end{pmatrix},\qquad n=1,2,\cdots\qquad (a_i,p_i\text{ は有理数})$$

の形の分布（これは可算無限個しかない）が $\mathcal{P}^1$ の中で稠密なことをいえばよい．

これには任意の分布が上の形の分布の列の弱位相極限(すなわち§2.5b)でのべた意味の極限)になることをいえばよい．この証明は，分布関数を考えると簡単にできるから，読者にまかせる．

$\mathcal{P}^1$ が弱位相(すなわち ρ 位相)において第2可算性をもつから，距離空間 $(\mathcal{P}^1, \rho)$ は可分距離空間である．しかしこの距離空間は完備ではない．そこで新しい距離 $d(\mu, \nu)$ を $\mathcal{P}^1$ の中に入れ，d 位相が ρ 位相(すなわち弱位相)と一致し，しかも $(\mathcal{P}^1, d)$ が完備可分距離空間にできるかという問題がおこる．実はそれは可能なので，Lévy の導入した距離(Lévy 距離)がこのようなものである．

$F(x)$ を μ の分布関数とし，$y=F(x)$ のグラフを考える．$F(x)$ の不連続点 a では $(a, F(a-))$ と $(a, F(a+))$ を線分でつないでおくと，このグラフは連続曲線となる．この曲線と直線 $x+y=t$ との交点(これはただ一つ)と，点 $(t, 0)$ との距

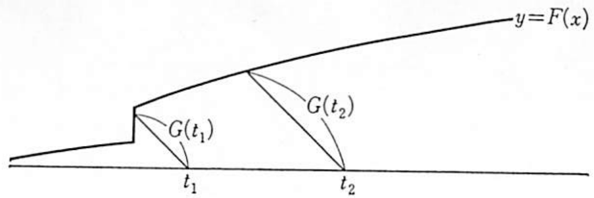


図 2.2

離を $G(t)$ とする．$G(t)$ は t の増加関数で

$$|G(t)-G(s)| \leqq |t-s|, \qquad \lim_{t\to-\infty} G(t) = 0, \qquad \lim_{t\to\infty} G(t) = \sqrt{2}.$$

逆にこの性質をもつ関数はただ一つの分布 μ に上の意味で対応する．このような関数 G の全体を $\mathcal{G}$ とする．$\mathcal{G}$ が距離

$$d(G_1, G_2) = \sup_t |G_1(t)-G_2(t)|$$

で完備可分距離空間となることは容易に証明できる．したがって分布 μ, ν の距離 $d_L(\mu, \nu)$ (**Lévy 距離**という)を

$$d_L(\mu, \nu) = d(G_\mu, G_\nu) \qquad (G_\mu, G_\nu \text{ は } \mu, \nu \text{ に対応する } G)$$

で定義すれば，$(\mathcal{P}, d_L)$ も完備可分距離空間となる．残る所は，d_L 位相が弱位相と一致することの証明である．

d_L 位相も弱位相も第1可算性をもつから，

$$\mu_n \longrightarrow \mu \ (弱) \iff d_L(\mu_n, \mu) \longrightarrow 0$$

をいえばよい．μ_n, μ に対応する分布関数，G 関数をそれぞれ $F_n, F; G_n, G$ とすると，上の同等関係は

$$F_n(a) \longrightarrow F(a) \ (F \text{ の連続点 } a \text{ で}) \iff \|G_n-G\|_\infty \longrightarrow 0$$

とかける．

$\|G_n-G\|_\infty \to 0$ とする．a を F の連続点とする．$(a, F_n(a))$ を通る直線 $x+y=a+F_n(a)$ と $y=F(x)$ のグラフとの交点を (a_n, b_n) とする．$a_n \geqq a$ のときには $F_n(a)-b_n \leqq \|G_n-G\|_\infty \to 0 \ (n\to\infty)$，したがって

$$\begin{aligned} 0 \leqq F_n(a)-F(a) &\leqq (F_n(a)-b_n)+(b_n-F(a)) \\ &\leqq \|G_n-G\|_\infty + (F(a_n)-F(a)) \longrightarrow 0. \end{aligned}$$

$a_n < a$ のときにも

$$\begin{aligned} 0 \leqq F(a)-F_n(a) &\leqq (F(a)-b_n)+(b_n-F_n(a)) \\ &\leqq (F(a)-F(a_n-)) + \|G_n-G\|_\infty \longrightarrow 0. \end{aligned}$$

これから

$$\|G_n-G\|_\infty \longrightarrow 0 \implies F_n(a) \longrightarrow F(a) \qquad (F \text{ の連続点 } a \text{ で})$$

がでる．逆をいうには，F の連続点 $a_0, a_1, \cdots, a_m$ を

$$F(a_0) < \varepsilon, \quad F(a_m) > 1-\varepsilon, \quad 0 < a_k - a_{k-1} < \varepsilon, \quad k = 1, 2, \cdots, m$$

をみたすようにとると，十分大きい n に対して

$$|F_n(a_k) - F(a_k)| < \varepsilon, \quad k = 0, 1, 2, \cdots, m$$

となる．$x+y=t$ と F_n, F のグラフとの交点をそれぞれ $(\xi_n, \eta_n), (\xi, \eta)$ とすると，$\xi < a_0$ または $\xi > a_m$ のときには，この2点の距離すなわち $|G_n(t)-G(t)|$ は $2\sqrt{2}\varepsilon$ より小さい．また $a_{k-1} < \xi < a_k$ のときには，この距離は $\sqrt{2}\varepsilon + 2\sqrt{2}\varepsilon$ よりも小さい．これは $y=F_n(x)$，$y=F(x)$ のグラフをかいてみると，容易にわかる(F_n, F が増加関数であることに注意せよ)．したがって

$$\|G_n-G\|_\infty < \sqrt{2}\varepsilon + 2\sqrt{2}\varepsilon$$

となり，目的は達せられた．

例題 2.7 (i) $d_L(T_{0,c}, \delta)$ を計算せよ．

(ii) "$\mathcal{M}$ $(\subset \mathcal{P}^1)$ の弱位相による閉包がコンパクトであること"，"$\mathcal{M}$ が d_L に

ついて全有界（任意の $\varepsilon>0$ に対し，有限個の ε 近傍（d_L について）でおおわれること）”，“$\liminf_{a\to\infty}\inf_{\mu\in\mathcal{M}}\mu[-a,a]=1$” の三つの条件は同等であることを示せ（定理2.11参照）．

§2.8　d 次元の分布

$\boldsymbol{R}^d$ の上の正則確率測度を **d 次元分布**という．§2.5, §2.6, §2.7の3節にわたって論じた1次元分布の性質はほとんどそのまま d 次元分布に対してもなりたつ．

始めに d_i 次元の分布 μ_i $(i=1,2,\cdots,n)$ の完備直積は $d\left(=\sum_i d_i\right)$ 次元の分布であることを注意しておく．したがって $\mu_1,\mu_2,\cdots,\mu_d$ がいずれも1次元の分布ならば

$$\mu=\overline{\mu_1\times\mu_2\times\cdots\times\mu_d}\qquad(\text{完備直積})$$

は d 次元の分布である．特に

$$\mu_i=\begin{pmatrix}a_{i1},\ a_{i2},\ \cdots\\ p_{i1},\ p_{i2},\ \cdots\end{pmatrix},\qquad i=1,2,\cdots,d$$

のときには，その完備直積 μ は

$$\mu=\begin{pmatrix}(a_{1j_1},\ a_{2j_2},\ \cdots,\ a_{dj_d})\\ p_{1j_1}\ p_{2j_2}\ \cdots\ p_{dj_d}\end{pmatrix}_{j_1,j_2,\cdots,j_d}$$

で与えられる．またすべての $i=1,2,\cdots,d$ に対し，μ_i が密度 f_i をもてば，μ は密度

$$f(x_1,x_2,\cdots,x_d)=f_1(x_1)f_2(x_2)\cdots f_d(x_d)$$

をもつ．また μ_i がすべて特異であれば，μ も特異である．任意の分布が純不連続分布，絶対連続分布，特異分布の凸結合となるという Lebesgue 分解定理は d 次元でもなりたつ．

絶対連続分布と密度をもつ分布とは一致する．密度をもつ分布のうち最も有名なものは **Gauss 分布** $N_{m,V}$ で，その密度は

$$f(x_1,x_2,\cdots,x_d)=\frac{1}{(2\pi)^{d/2}\sqrt{\det V}}\exp\left\{-\frac{1}{2}{}^t(x-m)V^{-1}(x-m)\right\}$$

で与えられる．ここに $m\in\boldsymbol{R}^d$ でありまた V は強正定対称実行列である．$x-m$ は列ベクトル（すなわち $d\times1$ 行列）であらわし，${}^t(x-m)$ はその転置行列（$1\times d$

行列）である．${}^t(x-m)V^{-1}(x-m)$ は内積の形に

$$(V^{-1}(x-m),\ (x-m))$$

とかいてもよい．f が分布の密度となっていることをいうには f の $\boldsymbol{R}^d$ 上の積分が1に等しいことをいえばよい．仮定により

$$V=UDU^{-1},\qquad U\text{ は直交行列，}D\text{ は対角行列}\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_d\end{bmatrix}$$

となるから，変数変換

$$x=m+Uy$$

を施すと，$|\det U|=1$ により

$$\begin{aligned}
&\int\cdots\int\exp\left\{-\frac{1}{2}{}^t(x-m)V^{-1}(x-m)\right\}dx_1\cdots dx_d\\
&=\int\cdots\int\exp\left\{-\frac{1}{2}{}^ty{}^tUV^{-1}Uy\right\}dy_1\cdots dy_d\\
&=\int\cdots\int\exp\left\{-\frac{1}{2}{}^tyU^{-1}V^{-1}Uy\right\}dy_1\cdots dy_d\qquad({}^tU=U^{-1}\text{ により})\\
&=\int\cdots\int\exp\left\{-\frac{1}{2}{}^ty(U^{-1}VU)^{-1}y\right\}dy_1\cdots dy_d\\
&=\int\cdots\int\exp\left\{-\frac{1}{2}{}^tyD^{-1}y\right\}dy_1\cdots dy_d\\
&=\int\cdots\int\exp\left\{-\sum_i\frac{y_i^2}{2\lambda_i}\right\}dy_1\cdots dy_d\\
&=\prod_i\int_{R^1}\exp\left\{-\frac{y_i^2}{2\lambda_i}\right\}dy_i=\prod_i\sqrt{2\pi\lambda_i}=(2\pi)^{d/2}\sqrt{\prod_i\lambda_i}\\
&=(2\pi)^{d/2}\sqrt{\det D}=(2\pi)^{d/2}\sqrt{\det U\det V\det U^{-1}}=(2\pi)^{d/2}\sqrt{\det V}.
\end{aligned}$$

これから直ちに f の $\boldsymbol{R}^d$ 上の積分が1に等しいことがわかる．

d 次元の純不連続分布の例として**多項分布**をあげておこう．これは

$$P_{p_1,p_2,\cdots,p_d,k}=\begin{pmatrix}(k_1,\ k_2,\ \cdots,\ k_d)\\ \dfrac{k!}{k_1!\,k_2!\cdots k_d!}p_1^{k_1}p_2^{k_2}\cdots p_d^{k_d}\end{pmatrix}_{k_i=1,2,\cdots,\ \sum_{i=1}^d k_i=k}$$

で与えられる．ここに $p_1,p_2,\cdots,p_d$ はその和が1に等しい正数であり，k は与えられた自然数である．$(k_1,k_2,\cdots,k_d)$ に付与された確率は

$$(p_1t_1+p_2t_2+\cdots+p_dt_d)^k$$

の展開における $t_1^{k_1}t_2^{k_2}\cdots t_d^{k_d}$ の係数に等しいことから，多項分布の名が与えられたのである．上の多項式で

$$t_1=t_2=\cdots=t_d=1$$

とおいて見ると，全確率が1に等しいことがわかる．

d 次元分布の全体を $\mathcal{P}^d$ であらわす．$\mu\in\mathcal{P}^d$ の分布関数 $F(x_1,x_2,\cdots,x_d)$ を

$$F(x_1,x_2,\cdots,x_d)=\mu((-\infty,x_1]\times(-\infty,x_2]\times\cdots\times(-\infty,x_d])$$

で定義し，$\mathcal{P}^d$ の中の位相，収束，たたみこみを1次元と同様に定義して，1次元の場合とほとんど同様の理論を展開できる．また $\mu\in\mathcal{P}^d$ の特性関数 $\varphi(z)$ を

$$\varphi(z)=\int_{R^d}e^{i(x,z)}\mu(dx),\qquad z\in \boldsymbol{R}^d\qquad ((x,z)\text{ は }x,z\text{ の内積})$$

で定義すれば，1次元の結果はそのまま d 次元に拡張される．

例題 2.8　(i)　Gauss 分布 $N_{m,V}$ の特性関数 $\varphi(z)$ は

$$\varphi(z)=\exp\left\{i(m,z)-\frac{1}{2}{}^tzVz\right\}$$

に等しいことを示せ．

［ヒント］　特性関数をあらわす積分の式で変換

$$x=m+Uy,\qquad U\text{ は }V\text{ を対角化する直交行列}$$

を施せ．

(ii)　$m\in\boldsymbol{R}^d$，V は正定対称実行列(必ずしも強正定ではない)とする．このときにも問題(i)の $\varphi(z)$ の式はやはりある d 次元分布の特性関数であることを示せ．(この分布も **Gauss 分布**といい，$N_{m,V}$ であらわす.)

［ヒント］　$V_n=V+n^{-1}I$ (I は単位行列) は強正定対称実行列であるから，$\mu_n=N_{m,V_n}$ は意味をもち，μ_n の特性関数 $\varphi_n(z)$ は

$$\varphi_n(z)=\exp\left\{i(m,z)-\frac{1}{2}{}^tzV_nz\right\}$$

$$\xrightarrow[(n\to\infty)]{}\varphi(z)=\exp\left\{i(m,z)-\frac{1}{2}{}^tzVz\right\}.$$

しかもこの収束は $z=0$ の近傍で一様であるから，Lévy の収束定理(の d 次元への拡張)が利用できる．

(iii)　$N_{m_1,V_1}*N_{m_2,V_2}=N_{m_1+m_2,V_1+V_2}$ を特性関数を用いて証明せよ．

(iv)　多項分布 $P_{p_1,p_2,\cdots,p_d,k}$ の特性関数は

$$(p_1e^{iz_1}+p_2e^{iz_2}+\cdots+p_de^{iz_d})^k$$

であることを示せ．

(v)　前問の結果を利用して

$$P_{p_1,p_2,\cdots,p_d,k}*P_{p_1,p_2,\cdots,p_d,r}=P_{p_1,p_2,\cdots,p_d,k+r}$$

を証明せよ．

§2.9　$\boldsymbol{R}^\infty$ の上の分布

すでに§2.5でのべたように，

$$\boldsymbol{R}^\infty=\boldsymbol{R}^1\times\boldsymbol{R}^1\times\cdots$$

は数列空間とよばれ，$\boldsymbol{R}^\infty$ の上の積位相は距離

$$\rho((x_n),(y_n))=\sum_{n=1}^{\infty}2^{-n}[|x_n-y_n|\wedge 1]$$

で与えられ，距離空間 $(\boldsymbol{R}^\infty,\rho)$ は完備かつ可分である．したがって $\boldsymbol{R}^\infty$ の上の正則確率測度(これを簡単に $\boldsymbol{R}^\infty$ の上の**分布**という)は標準である．また $\boldsymbol{R}^\infty$ の Borel 集合族 $\mathcal{B}(\boldsymbol{R}^\infty)$ ($\mathcal{B}^\infty$ とかく)は

$$\mathcal{B}^1\times\mathcal{B}^1\times\cdots\qquad(\text{積 }\sigma\text{ 加法族})$$

と一致する．前節では注意しなかったが，これらの事実は $\boldsymbol{R}^d$ ($d=1,2,\cdots$) の上でももちろんなりたつ．

$\boldsymbol{R}^\infty$ の点 $(x_1,x_2,\cdots)$ に $\boldsymbol{R}^n$ の点 $(x_1,x_2,\cdots,x_n)$ を対応させることにより，写像 $p_n:\boldsymbol{R}^\infty\to\boldsymbol{R}^n$ が定義される．同様にして $n<m$ のとき

$$p_{n,m}:\ \boldsymbol{R}^m\longrightarrow\boldsymbol{R}^n,\qquad (x_1,x_2,\cdots,x_m)\longmapsto(x_1,x_2,\cdots,x_n)$$

が定義される．明らかに $p_n, p_{n,m}$ は連続(したがって Borel 可測)である．また

$$p_{n,m}\circ p_{m,l}=p_{n,l},$$
$$p_{n,m}\circ p_m=p_n$$

も明らかである．

μ を $\boldsymbol{R}^\infty$ の上の分布とする．$\boldsymbol{R}^\infty$ も $\boldsymbol{R}^n$ ($n<\infty$) も完備可分距離空間であるから，像測度

$$\mu_n=\mu p_n^{-1}$$

は補題2.3(§2.4)により $\boldsymbol{R}^n$ の上の分布である．このようにして得られる分布列

$\{\mu_n\}$ の間には

(K)　$$\mu_n = \mu_m p_{n,m}^{-1}, \qquad n < m$$

がなりたつことは

$$p_n^{-1}(E) = (p_{n,m}\circ p_m)^{-1}(E) = p_m^{-1}(p_{n,m}^{-1}(E))$$

から明らかである．上の関係 (K) を Kolmogorov の**両立条件**という．この逆として次の重要な定理がなりたつ．

定理 2.20 (Kolmogorov の拡張定理)　分布の列

$$\mu_n \in \mathcal{P}^n, \qquad n = 1, 2, \cdots \qquad (\mathcal{P}^n \text{ は } n \text{ 次元分布の全体})$$

が Kolmogorov の両立条件をみたすならば，$\boldsymbol{R}^\infty$ の上に一つしかもただ一つの分布 μ が存在して

$$\mu_n = \mu p_n^{-1}, \qquad n = 1, 2, \cdots.$$

証明　$\boldsymbol{R}^\infty$ の上の集合族

$$\mathcal{B}_n = \{p_n^{-1}(E_n) \mid E_n \in \mathcal{B}^n\}, \qquad n = 1, 2, \cdots,$$

$$\mathcal{A} = \bigcup_n \mathcal{B}_n \qquad (\textstyle\bigcup \text{ は集合和を示す})$$

を考えよう．誤解のないように断っておくが，$\mathcal{A}$ は少なくとも一つの $\mathcal{B}_n$ に属する集合の全体である．明らかに $\mathcal{B}_n$ は $\boldsymbol{R}^\infty$ の上の σ 加法族であるが，$\mathcal{A}$ は $\boldsymbol{R}^\infty$ の上の加法族である．また p_n の定義から

$$p_n(p_n^{-1}(E_n)) = E_n$$

であるから，$\mathcal{B}_n$ の元 B を $B = p_n^{-1}(E_n)$ $(E_n \in \mathcal{B}^n)$ であらわすとき，この E_n は B で一意に定まる．$m>n$ のときには $\mathcal{B}_n$ の元 $B = p_n^{-1}(E_n)$ は

$$B = (p_{n,m}\circ p_m)^{-1}(E_n) = p_m^{-1}(p_{n,m}^{-1}(E_n))$$

ともかけるから，B は $\mathcal{B}_m$ に属する．したがって

$$\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{A} \subset \mathcal{B}^\infty$$

である．$\mathcal{B}^\infty = \mathcal{B}^1 \times \mathcal{B}^1 \times \cdots$ により

$$\mathcal{B}^\infty = \sigma[\mathcal{A}]$$

であることがわかる．

さて $B \in \mathcal{B}_n$ は $B = p_n^{-1}(E_n)\,(E_n \in \mathcal{B}^n)$ と一意にかけるから

$$\tilde{\mu}_n(B) = \mu_n(E_n)$$

と定義すれば，$\tilde{\mu}_n$ が $\mathcal{B}_n$ を定義域とする $\boldsymbol{R}^\infty$ 上の確率測度となる．$\mathcal{B}_n$ の元 $B = p_n^{-1}(E_n)\,(E_n \in \mathcal{B}^n)$ は

$$B = p_m^{-1}(p_{n,m}^{-1}(E_n)), \qquad m > n$$

ともかけるから，

$$\tilde{\mu}_m(B) = \mu_n(p_{n,m}^{-1}(E_n))$$

であるが，右辺は両立条件の仮定により $\mu_m(E_n)$ すなわち $\tilde{\mu}_n(B)$ に等しい．これは $m>n$ のとき $\tilde{\mu}_m$ が $\tilde{\mu}_n$ の拡張となっていることを意味する．したがって $B \in \mathcal{A} = \bigcup_n \mathcal{B}_n$ に対し

$$\tilde{\mu}(B) = \tilde{\mu}_n(B), \qquad B \in \mathcal{B}_n$$

と定義すれば，$\tilde{\mu}(B)$, $B \in \mathcal{A}$, が確定する．$\tilde{\mu}$ は初等確率測度であることも明らかである．

さてもし μ, μ' が定理の条件をみたす $\boldsymbol{R}^\infty$ 上の分布ならば，μ, μ' はともに $\mathcal{A}$ の上では $\tilde{\mu}$ と一致する．$\mathcal{A}$ は $\mathcal{B}^\infty$ を生成する加法族であるから，確率測度一致の定理により $\mathcal{B}^\infty$ の上で μ, μ' は一致するが，μ, μ' は正則であるから，結局完全に一致する．これで一意性は証明された．

残るのは存在証明である．上の考察から定理の条件をみたす分布 μ は上に構成した初等確率測度 $\tilde{\mu}$ の拡張であるべきであるから，$\tilde{\mu}$ が確率測度拡張定理の条件

$$\text{“}A_n \in \mathcal{A}\ (n=1, 2, \cdots),\ A_1 \supset A_2 \supset \cdots,\ a = \inf \tilde{\mu}(A_n) > 0\text{”}$$
$$\Longrightarrow \bigcap_n A_n \neq \phi$$

をみたすことをいえばよい．$A_n \in \mathcal{A}$ から $A_n \in \mathcal{B}_{k(n)}$ であるが，A_n と A_{n+1} との間に A_n をいくつかいれることにより

$$A_n \in \mathcal{B}_n, \qquad n = 1, 2, \cdots$$

として一般性を失わない．したがって

$$A_n = p_n^{-1}(E_n), \qquad E_n \in \mathcal{B}^n, \qquad \tilde{\mu}(A_n) = \mu_n(E_n), \qquad n = 1, 2, \cdots$$

となる．μ_n は K 正則であるから，コンパクト集合 $K_n \subset E_n$ が存在して

$$\mu_n(E_n - K_n) < 2^{-n-1}a, \qquad n = 1, 2, \cdots$$

となる．$C_n = p_n^{-1}(K_n)$ とおくと，

$$\tilde{\mu}(C_n) = \mu_n(K_n), \qquad \tilde{\mu}(A_n - C_n) = \mu_n(E_n - K_n) < 2^{-n-1}a.$$

$D_n = \bigcap_{i=1}^n C_i$ とおくと，$\{A_n\}$ は減少列であるから

$$\tilde{\mu}(A_n - D_n) = \tilde{\mu}\left(\bigcap_{i=1}^n A_i - \bigcap_{i=1}^n C_i\right) \leqq \sum_{i=1}^n \tilde{\mu}(A_i - C_i)$$

$$< \sum_{i=1}^{n} 2^{-i-1} a < \frac{a}{2},$$

故に

$$\tilde{\mu}(D_n) = \tilde{\mu}(A_n) - \tilde{\mu}(A_n - D_n) \geqq a - \frac{a}{2} > 0.$$

これから，D_n が空でないことがわかる．D_n の中から

$$\xi_n = (x_n{}^1, x_n{}^2, \cdots) \qquad \text{(便宜上座標番号を肩につける)}$$

をとると，すべての n に対し

$$\xi_n \in D_n \subset D_1 \subset C_1,$$

故に

$$x_n{}^1 = p_1(\xi_n) \in K_1, \qquad n = 1, 2, \cdots.$$

K_1 がコンパクトであるから，$\{x_n{}^1\}$ は収束部分列

$$x_{k(1,1)}{}^1,\ x_{k(1,2)}{}^1,\ x_{k(1,3)}{}^1,\ \cdots$$

をもつ．

$$\xi_{k(1,n)} \in D_{k(1,n)} \subset D_2 \subset C_2, \qquad n = 2, 3, 4, \cdots$$

であるから，

$$(x_{k(1,n)}{}^1, x_{k(1,n)}{}^2) = p_2(\xi_{k(1,n)}) \in K_2, \qquad n = 1, 2, \cdots.$$

故にこの2次元の点の列は収束部分列

$$(x_{k(2,n)}{}^1, x_{k(2,n)}{}^2), \qquad n = 1, 2, \cdots$$

をもつ．同様にくりかえして収束点列

$$(x_{k(m,n)}{}^1, x_{k(m,n)}{}^2, \cdots, x_{k(m,n)}{}^m), \qquad n = 1, 2, \cdots$$

がすべての m に対して得られる．しかも $m' > m$ のときには添字の列 $k(m', 1)$, $k(m', 2), \cdots$ は $k(m, 1), k(m, 2), \cdots$ の部分列となっている．したがって

$$\xi_{k(n,n)} = (\tilde{x}_n{}^1, \tilde{x}_n{}^2, \cdots)$$

とおくと，おのおのの m に対し

$$(\tilde{x}_n{}^1, \tilde{x}_n{}^2, \cdots, \tilde{x}_n{}^m), \qquad n = 1, 2, \cdots$$

m 項以後は，上の m 次元の収束点列の部分列であり，したがって収束点列である(対角線論法!)．故に

$$x^m = \lim_{n\to\infty} \tilde{x}_n{}^m, \qquad m = 1, 2, \cdots$$

が定まり，

$$(\tilde{x}_n{}^1, \tilde{x}_n{}^2, \cdots, \tilde{x}_n{}^m) \longrightarrow (x^1, x^2, \cdots, x^m) \quad (n \to \infty), \qquad m = 1, 2, \cdots$$

となる．$n \geqq m$ である限り

$$(\tilde{x}_n{}^1, \tilde{x}_n{}^2, \cdots, \tilde{x}_n{}^m) = p_m(\xi_{k(n,n)}) \in p_m(D_{k(n,n)}) \subset p_m(D_m) \subset p_m(C_m) = K_m,$$

K_m はコンパクトであるから

$$(x^1, x^2, \cdots, x^m) \in K_m, \qquad m = 1, 2, \cdots.$$

これは

$$\xi = (x^1, x^2, \cdots) \in p_m{}^{-1}(K_m) = C_m, \qquad m = 1, 2, \cdots$$

を意味する．故に

$$\xi \in \bigcap_m C_m \subset \bigcap_m A_m \quad \text{すなわち} \quad \bigcap_m A_m \neq \phi. \qquad \blacksquare$$

例題 2.9 Kolmogorov の拡張定理を用いて1次元の分布の列 $\{\mu_n\}$ の直積 $\mu = \mu_1 \times \mu_2 \times \cdots$ を定義せよ．

［ヒント］ $\nu_n = \mu_1 \times \mu_2 \times \cdots \times \mu_n \in \mathcal{P}^n$ $(n = 1, 2, \cdots)$ が Kolmogorov の両立条件をみたすことをいえばよい．

第3章　確率論の基礎概念

本章の目的は，一般試行に対する確率論の基礎概念を測度論の言葉で厳密に定義することにある．基礎におく確率空間は常に (Ω, P) であらわし，Ω の一般元を ω であらわす．直観的には Ω は考察する試行の見本空間であり，ω は出現する見本点一般をあらわし，$P(A)$ は ω が A に入る確率 (すなわち A の中の見本点が出現する確率) をあらわす．本講では理論を透明にするため，(Ω, P) **を可分完全確率空間である**と仮定しておく．(可分性，完全性の定義は §3.1 でのべる.) 実際，応用上興味のある問題に必要な確率空間はすべて可分完全である．

確率空間を基礎として確率論を測度論的に構築するという考えは E. Borel の大数の強法則の証明 (1909)，N. Wiener の Brown 運動の研究 (1920-24) にもあらわれているが，確率論全分野をこの考えで貫くことができることを示したのは A. Kolmogorov の "確率論の基礎概念" (1933) であろう．そこでは Kolmogorov は基礎におく確率空間に完備性だけを要求した．しかし Kolmogorov 自身が後になって '確率変数の確率法則の定義' に関連して，(Ω, P) に完全性の仮定をおくことを提唱した．しかし本講ではさらに可分性の仮定を追加することにした．なお多少の差違はあるが，(Ω, P) を制限する類似の提案が G. Mackey, D. Blackwell, L. Schwartz, P. Cartier などによってなされている．

§3.1　可分完全確率測度

μ を S の上の完備確率測度とする．任意の μ 可測写像 $f: S \to \boldsymbol{R}^1$ に対し，像測度 μf^{-1} が正則となるとき，μ を**完全** (perfect) であるという．このとき確率空間 (S, μ) も完全であるという．例えば，完備可分距離空間の上の正則確率測度は，補題 2.3 により完全である．

S の上の集合族 $\mathscr{A}$ が**分離族**である (または $\mathscr{A}$ が S の2点を**分離**する) とは，任意の $s_1, s_2 \in S$ $(s_1 \neq s_2)$ に対し，$A \in \mathscr{A}$ が存在して

$$1_A(s_1) \neq 1_A(s_2)$$

となることである．$\mathscr{A}$ が分離族ならば，$\mathscr{A}$ で生成される σ 加法族 $\sigma[\mathscr{A}]$ は当然分離族であるが，逆もなりたつ．もし $\sigma[\mathscr{A}]$ が分離族であるにもかかわらず，$\mathscr{A}$ が分離族でなければ，$s_1, s_2 \in S$ $(s_1 \neq s_2)$ が存在して，すべての $A \in \mathscr{A}$ に対し

$$1_A(s_1) = 1_A(s_2)$$

となる．この等式のなりたつ A の全体 $\mathscr{B}$ は，

$$1_S = 1, \quad 1_{A^c} = 1-1_A, \quad 1_{\cap A_n} = \prod 1_{A_n}$$

により，S の上の σ 加法族である．しかも上の仮定から $\mathscr{B}$ は $\mathscr{A}$ を含む．したがって $\mathscr{B} \supset \sigma[\mathscr{A}]$．これは $\sigma[\mathscr{A}]$ が分離族でないことを示し，矛盾である．

μ を S の上の完備確率測度とする．$\mathscr{D}(\mu)$ が可算分離族を含むとき，μ を**可分**という．このとき確率空間 (S, μ) も可分であるという．

補題 3.1　$T \subset \boldsymbol{R}^1$, μ は T の上の完備確率測度，$i: T \to \boldsymbol{R}^1$ は恒等写像，$\nu = \mu i^{-1}$ とするならば

$$\mu \text{ が } K \text{ 正則} \iff \nu \text{ が正則}.$$

証明　ν が $\boldsymbol{R}^1$ の上の完備確率測度であるから，ν の正則性と K 正則性とは一致する．$T, \boldsymbol{R}^1$ はともに Hausdorff 空間であるから，そこではコンパクト集合は閉集合であり，$\theta = \mu$ (または ν) の K 正則性は "任意の $A \in \mathscr{D}(\theta)$ に対して，コンパクト集合列 $K_n \in \mathscr{D}(\theta)$ $(n=1, 2, \cdots)$ を $K_n \subset A$, $\theta(K_n) \to \theta(A)$ をみたすようにとれる" ことと同等である．$\nu = \mu i^{-1}$, $i^{-1}(A) = A \cap T$ により，

$$T \in \mathscr{D}(\nu), \quad \nu(T) = 1, \quad \mu = \nu|_T$$

がでる．これだけ注意すれば，μ の K 正則性と ν の正則性との同値なことは，容易に証明される．∎

定理 3.1　可分完全確率測度は $\boldsymbol{R}^1$ のある部分集合の上の K 正則確率測度と強同型である．逆もまた真である．

証明　μ を S の上の可分完全確率測度とすると，$\mathscr{D}(\mu)$ は可算分離族 $\mathscr{A} = \{A_1, A_2, \cdots\}$ を含む．写像 $f: S \to \boldsymbol{R}^1$ を

$$f(s) = \sum_{n=1}^{\infty} 3^{-n} e_n(s), \quad e_n = 1_{A_n}, \ n = 1, 2, \cdots$$

で定義する．$A_n \in \mathscr{D}(\mu)$ により，$e_n: S \to \boldsymbol{R}^1$ は μ 可測であるから，$f: S \to \boldsymbol{R}^1$ も μ 可測である．μ は完全であるから，$\nu = \mu f^{-1}$ は $\boldsymbol{R}^1$ の上の正則確率測度である．f の定義により，f が単射であることは容易にわかる．したがって f の値域をその像 $T = f(S)$ $(\subset \boldsymbol{R}^1)$ に制限して得られる写像 $g: S \to T$ は全単射である．

$$\theta = \mu g^{-1}$$

とおくと，g が全単射であるから，θ は μ と (g に関して) 強同型である．$i: T \to \boldsymbol{R}^1$ を恒等写像とすると，

$$f = i \circ g,$$

$$\nu = \mu f^{-1} = (\mu g^{-1}) i^{-1} = \theta i^{-1}.$$

ν が $\boldsymbol{R}^1$ の上の正則確率測度であるから，上の式から θ の K 正則性がでる (補題 3.1)．

これで定理の前半は証明された．

後半を証明するには，

$$S \subset \boldsymbol{R}^1, \quad \mu \text{ は } K \text{ 正則}$$

として一般性を失わない．

$$\mathscr{A} = \{(-\infty, r) \cap S \,|\, r \in \boldsymbol{Q}\}$$

は S の上の可算分離族で $\mathscr{D}(\mu)$ に含まれるから，μ は可分である．μ の完全性を示すため，任意の μ 可測な $f: S \to \boldsymbol{R}^1$ を考えよう．$i: S \to \boldsymbol{R}^1$ を恒等写像とし

$$\mu_1 = \mu i^{-1}$$

とおくと，μ の K 正則性から μ_1 が $\boldsymbol{R}^1$ の上の正則確率測度であることがわかる (補題 3.1)．$f: S \to \boldsymbol{R}^1$ を任意に拡張して

$$f_1: \boldsymbol{R}^1 \longrightarrow \boldsymbol{R}^1$$

を定義すると，

$$\mu_1(\boldsymbol{R}^1 - S) = \mu(i^{-1}(\boldsymbol{R}^1 - S)) = \mu(\phi) = 0$$

により，$f_1: \boldsymbol{R}^1 \to \boldsymbol{R}^1$ が μ_1 可測で

$$\mu_1 f_1^{-1} = \mu f^{-1}$$

となることが容易にわかる．$\boldsymbol{R}^1$ は完備可分距離空間であるから，すでに注意したように，その上の正則確率測度 μ_1 は完全である．したがって $\mu_1 f_1^{-1}$ すなわち μf^{-1} は正則となり，結局 μ も完全となる．∎

定理 3.1 に補充してつぎの事実を付加しておこう．

定理 3.1*　μ を S の上の可分完全確率測度，$\mathscr{A} = \{A_1, A_2, \cdots\} \subset \mathscr{D}(\mu)$ を S の上の分離族とすれば，μ は $\mu|_{\sigma[\mathscr{A}]}$ の Lebesgue 拡大である．

証明　前定理の証明に用いた $f: S \to \boldsymbol{R}^1$ は可測 $\sigma[\mathscr{A}]/\mathscr{B}^1$ であることは，1_{A_n}:

$S \to \boldsymbol{R}^1$ が ($A_n \in \sigma[\mathscr{A}]$ により) 可測 $\sigma[\mathscr{A}]/\mathcal{B}^1$ であることからすぐにわかる．ゆえに

$$\sigma[\mathscr{A}] \supset f^{-1}(\mathcal{B}^1).$$

また

$$e_n(s) = \varphi_n(f(s)), \qquad \varphi_n(x) = [3^n x] - 3[3^{n-1}x] \quad ([\ \ ] \text{は整数部分})$$

で，$\varphi_n : \boldsymbol{R}^1 \to \boldsymbol{R}^1$ は Borel 可測であるから，$B_n \equiv \varphi_n{}^{-1}(1) \in \mathcal{B}^1$,

$$A_n = e_n{}^{-1}(1) = \{s \mid f(s) \in B_n\} = f^{-1}(B_n) \in f^{-1}(\mathcal{B}^1), \qquad n = 1, 2, \cdots,$$

これから

$$\sigma[\mathscr{A}] \subset f^{-1}(\mathcal{B}^1), \qquad \text{したがって} \quad \sigma[\mathscr{A}] = f^{-1}(\mathcal{B}^1)$$

が得られる．$\nu = \mu f^{-1}$ は正則であるから，ν は $\nu|_{\mathcal{B}^1}$ の Lebesgue 拡大であり，したがって μ は $\mu|_{\sigma[\mathscr{A}]}$ の Lebesgue 拡大となる．∎

次の定理は**完全性が像測度に遺伝する**ことを示している．

定理 3.2 μ が S の上の完全確率測度，$f : S \to T$ とすると，$\nu = \mu f^{-1}$ は T の上の完全確率測度である．

証明 $g : T \to \boldsymbol{R}^1$ を ν 可測とすると，$h = g \circ f : S \to \boldsymbol{R}^1$ は μ 可測である．(なぜならば，$\nu = \mu f^{-1}$ により

$$\Gamma \in \mathcal{B}^1 \implies g^{-1}(\Gamma) \in \mathcal{D}(\nu) \implies h^{-1}(\Gamma) = f^{-1}(g^{-1}(\Gamma)) \in \mathcal{D}(\mu)$$

となるからである．) したがって μ の完全性により，μh^{-1} は正則．

$$\mu h^{-1} = (\mu f^{-1}) g^{-1} = \nu g^{-1}$$

により，νg^{-1} は正則．これは ν が完全であることを意味する．∎

定理 3.3 S が位相空間で，$\mathcal{B}(S)$ が可算族で生成される (例えば S が可算開基をもつ) とし，μ が S の上の完全確率測度とする．もし $\mathcal{D}(\mu)$ が $\mathcal{B}(S)$ を含めば，μ は可分かつ正則である．

証明 $\mathcal{B}(S)$ を生成する可算族 $\mathscr{A} = \{A_n\}$ をとる．1 点集合は閉集合であるから $\mathcal{B}(S)$ に属する．ゆえに $\mathcal{B}(S)$ は S の上の分離族である．したがって $\mathscr{A}$ も分離族である．$\mathscr{A} \subset \mathcal{D}(\mu)$ であるから，μ は可分かつ完全であり，定理 3.1* により，μ は $\mu|_{\sigma[\mathscr{A}]}$ (すなわち $\mu|_{\mathcal{B}(S)}$) の Lebesgue 拡大と一致する．これは μ が正則であることを意味する．∎

例題 3.1 (i) 完備可分距離空間 T の上の正則確率測度 μ は可分完全であることを示せ．

[ヒント] T は可算開基をもつから，μ は可分．完全性は補題 2.3 からでる．

(ii) μ が完備可分距離空間 T の上の完備確率測度で，$\mathcal{D}(\mu) \supset \mathcal{B}(T)$ とする．このとき μ の完全性と正則性とは一致することを示せ．

[ヒント] 完全$\Longrightarrow$正則，は定理 3.3 による．逆は補題 2.3 による．

(iii) (補題 2.3 の拡張) μ が S の上の完全確率測度，T が位相空間で，$\mathcal{B}(T)$ が可算族で生成され，写像 $f : S \to T$ が μ 可測 (すなわち可測 $\mathcal{D}(\mu)/\mathcal{B}(T)$) とするならば，$\nu = \mu f^{-1}$ は T の上の可分，完全かつ正則な確率測度となることを示せ．

[ヒント] f の μ 可測性により，$\mathcal{D}(\nu) \supset \mathcal{B}(T)$．定理 3.2 により，$\nu$ は完全である．ゆえに定理 3.3 により ν は可分かつ正則．

(iv) 標準確率測度は完全であることを示せ．

[ヒント] 空間から零集合をとり去れば，標準確率測度は可分完全となる (定義と定理 3.1 の逆の部分)．

(v) 濃度 $\boldsymbol{c}$ 以下の標準確率空間と可分完全確率空間とは一致することを示せ．

[ヒント] 前者が可分であることの証明に工夫を要する．

§3.2 事象と確率変数

可分完全確率空間 (Ω, P) を基礎において話を進める．$\alpha = \alpha(\omega)$ を Ω の一般元 ω に関する条件とする．$\alpha(\omega)$ を成立させる ω の値の全体 $\{\omega \mid \alpha(\omega)\}$ を α の**外延**といい $\{\alpha(\omega)\}$ または $\{\alpha\}$ であらわす．$\{\alpha\}$ が P 可測のとき，α を**事象**といい，$P\{\alpha\}$ を α のおこる**確率**という．$P\{\alpha\} = 1$ のとき，α が**ほとんど確実に** (almost surely) おこるといい，

$$\alpha(\omega) \quad \text{a.s.}$$

とかく．

事象の演算 $\alpha \vee \beta$, $\alpha \wedge \beta$, $\alpha^{\neg}$ は第 1 章の有限試行の場合と同様に定義される．この際 $\{\alpha\}$, $\{\beta\}$ が P 可測のときには

$$\{\alpha \vee \beta\} = \{\alpha\} \cup \{\beta\}, \qquad \{\alpha \wedge \beta\} = \{\alpha\} \cap \{\beta\}, \qquad \{\alpha^{\neg}\} = \{\alpha\}^c$$

も P 可測であることに注意すべきである．可算演算

$$\alpha_1 \vee \alpha_2 \vee \cdots \qquad \left(\bigvee_n \alpha_n \text{ または } (\exists n)\alpha_n \text{ ともかく}\right),$$

$$\alpha_1 \wedge \alpha_2 \wedge \cdots \qquad \left(\bigwedge_n \alpha_n \text{ または } (\forall n)\alpha_n \text{ ともかく}\right)$$

も同様に考えられる．$\{\alpha_n\} \in \mathcal{D}(P)$ $(n=1, 2, \cdots)$ ならば

$$\left\{\bigvee_n \alpha_n\right\} = \bigcup_n \{\alpha_n\} \in \mathcal{D}(P), \qquad \left\{\bigwedge_n \alpha_n\right\} = \bigcap_n \{\alpha_n\} \in \mathcal{D}(P).$$

"$\alpha_1(\omega), \alpha_2(\omega), \cdots$ の中の無限個がなりたつ" ことを

$$\alpha_n \quad \text{i.o.} \qquad (\text{i.o.}=\text{infinitely often})$$

であらわす．$\{\alpha_n\} \in \mathcal{D}(P)$ $(n=1, 2, \cdots)$ ならば

$$\{\alpha_n \text{ i.o.}\} = \left\{\bigwedge_n \bigvee_{k>n} \alpha_k\right\} = \bigcap_n \bigcup_{k>n} \{\alpha_k\} \in \mathcal{D}(P).$$

"$\alpha_1(\omega), \alpha_2(\omega), \cdots$ が有限個の例外を除いてすべてなりたつ" ことを

$$\alpha_n \quad \text{f.e.} \qquad (\text{f.e.}=\text{with a finite number of exceptions})$$

であらわす．$\{\alpha_n\} \in \mathcal{D}(P)$ $(n=1, 2, \cdots)$ ならば

$$\{\alpha_n \text{ f.e.}\} = \left\{\bigvee_n \bigwedge_{k>n} \alpha_k\right\} = \bigcup_n \bigcap_{k>n} \{\alpha_k\} \in \mathcal{D}(P).$$

$\{\alpha_n \text{ i.o.}\}$, $\{\alpha_n \text{ f.e.}\}$ はそれぞれ集合列 $\{\alpha_n\}$ の上，下極限に等しい．

事象の確率の性質は集合の P 測度の性質に帰着されるから，とりたてて列挙する必要もなかろう．ただつぎの定理は測度論の本にはあまり見られないが，確率論では極めて有用である．

定理 3.4 (Borel-Cantelli の補題)

$$\sum_{n=1}^{\infty} P\{\alpha_n\} < \infty \Longrightarrow P(\alpha_n \text{ i.o.}) = 0, \quad P(\alpha_n{}^{\neg} \text{ f.e.}) = 1.$$

証明 $A_n=\{\alpha_n\}$ とおくと，定理は

$$\sum_{n=1}^{\infty} P(A_n) < \infty \Longrightarrow P\left(\limsup_{n\to\infty} A_n\right) = 0, \quad P\left(\liminf_{n\to\infty} A_n{}^c\right) = 1$$

とかかれる．右辺の上，下極限をそれぞれ S, L とかくと，$L=S^c$．すべての n に対し

$$P(S) \leqq P\left(\bigcup_{k>n} A_k\right) \leqq \sum_{k>n} P(A_k).$$

仮定により，右辺は $n\to\infty$ のとき 0 に収束するから，

$$P(S) = 0, \qquad \text{したがって} \quad P(L) = P(S^c) = 1. \quad \blacksquare$$

事象についてはこの辺でとどめておいて，つぎに確率変数に進もう．まず最も簡単な実確率変数から始める．$X(\omega)$ を Ω の上で定義された実数値関数とする．実数 a に対し

$$X(\omega) \leqq a$$

は ω に関する条件であるが，これが事象と見なされてその確率が考えられるためには，その外延

$$\{X(\omega) \leqq a\} = X^{-1}(-\infty, a]$$

が P 可測でなければならない．測度論の言葉でいえば，$X(\omega)$ が P 可測実関数でなければならない．この考察を念頭において，(Ω, P) の上の P 可測実関数を**実確率変数**とよぶことにする．有限試行の場合には P 可測性は自動的になりたつので，わざわざことわる必要がなかったのである．

$X(\omega)$ が (Ω, P) の上の実確率変数とすれば，任意の Borel 集合 E に対し，

$$X^{-1}(E) \in \mathcal{D}(P) \quad \text{すなわち} \quad \{X(\omega) \in E\} \in \mathcal{D}(P)$$

となるから，'$X(\omega) \in E$' も事象となり，その確率

$$P\{X(\omega) \in E\} = P(X^{-1}(E))$$

が考えられる．E が Borel 集合でなくても

$$X^{-1}(E) \in \mathcal{D}(P)$$

であるならば，$P\{X(\omega) \in E\}=P(X^{-1}(E))$ が考えられる．しかも $P(X^{-1}(E))$ は $E\subset \boldsymbol{R}^1$ の関数として確率測度になっているから，これを $X(\omega)$ の**確率法則**といい，$P^X(E)$ であらわす．すなわち

$$\mathcal{D}(P^X) = \{E \mid X^{-1}(E) \in \mathcal{D}(P)\},$$
$$P^X(E) = P(X^{-1}(E)).$$

これから $X(\omega)$ の確率法則 P^X は像測度 PX^{-1} にほかならないことがわかる．

$X(\omega)$ に関する条件はつねに '$X(\omega) \in E$' の形にかけるから，P^X を知ることによって，$X(\omega)$ に関する条件のなりたつ確率について完全な情報を得ることができる．

$X(\omega)$ は P 可測であるから，E が Borel 集合のときには，$X^{-1}(E) \in \mathcal{D}(P)$ すなわち $E \in \mathcal{D}(P^X)$ となる．換言すれば

$$\mathcal{B}^1 \subset \mathcal{D}(P^X).$$

これだけでは $P^X \equiv PX^{-1}$ が $\boldsymbol{R}^1$ の上の正則確率測度（第2章の用語では1次元分

布）であるとはいえない．ただいえることは P^X がある1次元分布の拡張になっているということである．ここで始めて P の完全性の仮定が生きてくるので，この仮定により $P^X \equiv PX^{-1}$ が正則（すなわち1次元分布）となるのである．Kolmogorov が完全性の仮定をおくことを提唱した所以はここにある．

実確率変数の概念を拡張して **n 次元ベクトル確率変数** $(n=1, 2, \cdots, \infty)$ を (Ω, P) の上の $\boldsymbol{R}^n$ 値 P 可測関数として定義することができる．$X : \Omega \to \boldsymbol{R}^n$ が P 可測であるとはいうまでもなく，任意の $E \in \mathcal{B}^n$ に対して

$$X^{-1}(E) \in \mathcal{D}(P)$$

のことである．これは

$$X^{-1}(\mathcal{B}^n) \subset \mathcal{D}(P) \quad \text{または} \quad X \in \mathcal{D}(P)/\mathcal{B}^n$$

といってもよい．$\boldsymbol{R}^n$ は可算開基をもつから，例題3.1 (iii) を適用して，X の確率法測 $P^X = PX^{-1}$ が正則な確率測度となる．

もっと一般に T を位相空間とし，$\mathcal{B}(T)$ が可算族で生成されるとする．（例えば T が可算開基をもつときはこの条件がみたされる．）このときにも **T 値確率変数** $X(\omega)$ を (Ω, P) の上の T 値 P 可測関数として定義すれば，その確率法則 $P^X = PX^{-1}$ は T の上の可分，完全かつ正則な確率測度となることは例題3.1 (iii) により明らかである．

さらに一般の位相空間の中の値をとる確率変数も同様に定義できるが，そのときにはもはや確率法則の正則性は期待できない．応用上そのような一般的な確率変数を考える必要はおこらないから，本講では考えない．

以上は位相空間の中の値をとる確率変数であるが，さらに抽象的な集合の中の値をとる確率変数を考えることができる．S を一般の集合とし，$X(\omega)$ を Ω の上の S 値関数とする．S の上には，位相空間の上の Borel 集合族に相当するものがないから，$X(\omega)$ の P 可測性というのは無意味である．これにかわるものとして，P 可分性という概念を導入しよう．S 値関数 $X(\omega)$ が **P 可分**であるとは，S の上の可算分離族 $\mathcal{A}$ が存在して

$$X^{-1}(\mathcal{A}) \subset \mathcal{D}(P) \qquad (X^{-1}(\mathcal{A}) = \{X^{-1}(A) \mid A \in \mathcal{A}\})$$

となることである．さて P 可分 S 値関数 $X(\omega)$ を **S 値確率変数**とよぶことにしよう．このときにも $X(\omega)$ の確率法則 P^X は上の場合と同様に像測度 PX^{-1} で定義されるが，$X(\omega)$ の P 可分性から

$$\mathcal{A} \subset \mathcal{D}(PX^{-1}) = \mathcal{D}(P^X)$$

により，P^X の可分性が出る．また定理3.2により P の完全性から P^X の完全性がでる．かくして P^X は P と同様に可分完全確率測度となる．

さて上に定義した位相空間 T ($\mathcal{B}(T)$ は可算族で生成される）の中の値をとる確率変数は，（T の位相構造を無視して）T を単なる集合と見たときの T 値確率変数となっている．しかし逆は必ずしも真ではない．後者のときには $\mathcal{A}$ は $\mathcal{D}(PX^{-1})$ に含まれる任意の可算分離族でよいが，前者の場合には $\mathcal{A}$ は $\mathcal{B}(S)$ を生成するように選ぶ必要があるからである．普通位相空間の中の値をとる確率変数は位相を考慮した前者の意味にとるが，とくにこのことを強調したいときには **T 値位相確率変数**という．実確率変数は $\boldsymbol{R}^1$ 値位相確率変数のことであり，n 次元ベクトル確率変数は $\boldsymbol{R}^n$ 値位相確率変数のことである．後者はしばしば **n 次元確率ベクトル**と略称される．

以上のことを綜合すると，T が位相空間（ことわらなくても $\mathcal{B}(T)$ が可算族で生成されることを仮定）のときには，T 値位相確率変数 $X(\omega)$ は P 可測 T 値関数であり，S が一般の集合のときには S 値確率変数 $Y(\omega)$ は P 可分 S 値関数である．また $X(\omega)$ の確率法則 $P^X = PX^{-1}$ は T の上の可分，完全かつ正則な確率測度であり，$Y(\omega)$ の確率法則 $P^Y = PY^{-1}$ は S の上の可分かつ完全な確率測度である．T 値位相確率変数は（T の位相を無視して考えられた）T 値確率変数の特別なものとみなされるから，これから確率変数についてのべることは，当然位相確率変数にもなりたつ．

$X(\omega)$ を S 値確率変数とする．$X(\omega)$ の P 可分性により，S の上に可算分離族 $\mathcal{A} = \{A_n\}$ が存在するから，S の濃度 $\# S$ は連続濃度 $\boldsymbol{c}$ 以下でなければならない．実際

$$\varphi(s) = (e_1(s), e_2(s), \cdots), \qquad e_n = 1_{A_n}, \quad n = 1, 2, \cdots$$

とおくと，φ は S から $I = \{0, 1\}^\infty$ の中への写像で，$\mathcal{A}$ の分離性により，φ は単射である．したがって

$$\# S = \# \varphi(S) \leqq \# I = \boldsymbol{c}.$$

このことから確率変数の値域となり得る集合（位相確率変数のときには位相空間）の濃度は $\boldsymbol{c}$ をこえることはできない．

確率変数の**値域の変更**についてのべよう．$X(\omega)$ を S 値確率変数とし，T を

$X(\Omega)$ を含む濃度 $\boldsymbol{c}$ 以下の任意の集合とすると，$X(\omega)$ を Ω の上の T 値関数と考えることができる．このように考えたとき，$X(\omega)$ を $X_T(\omega)$ とかくことにする．たとえば整数値関数を実数値関数の特別のものと考えるなど，このような眺め方の変更は，解析学一般において見られることである．さて $X_T(\omega)$ は T 値確率変数であることを示そう．まず $X(\omega)$ は S 値確率変数であるから，S の上の可算分離族 $\{A_n\}$ があって

$$X^{-1}(A_n) \in \mathcal{D}(P), \qquad n = 1, 2, \cdots.$$

$B_n = A_n \cap U$ $(U = X(\Omega))$ とすれば，$X_T(\Omega) = X(\Omega) = U$ であるから，

$$X_T{}^{-1}(B_n) = X^{-1}(B_n) = X^{-1}(A_n) \in \mathcal{D}(P), \qquad n = 1, 2, \cdots.$$

$\{A_n\}$ が S の上の分離族であるから，$\{B_n\}$ は U の上の分離族である．また $\#(T-U) \leqq \#T \leqq \boldsymbol{c}$ であるから，$T-U$ は $\boldsymbol{R}^1$ のある部分集合と1対1に対応し，$T-U$ の上の可算分離族 $\{C_n\}$ がある．しかも $X_T(\Omega) = U$ により

$$X_T{}^{-1}(C_n) = \phi \in \mathcal{D}(P), \qquad n = 1, 2, \cdots.$$

$\{B_n, U_n \mid n = 1, 2, \cdots\}$ は $T = U + (T-U)$ の上の可算分離族である．したがって $X_T(\omega)$ は P 可分となり，T 値確率変数である．X の確率法則 $P^X = PX^{-1}$ は S の上の可分完全確率測度であるが，X_T の確率法則 $P^{X_T} = PX_T{}^{-1}$ は T の上の可分完全確率測度である．明らかに

$$P^X(U) = P^{X_T}(U) = P(\Omega) = 1.$$

位相確率変数についても同様のことがいえる．しかし $\mathcal{B}(S), \mathcal{B}(T)$ が可算族で生成され，$\mathcal{B}(T) \cap U = \mathcal{B}(S) \cap U$ を仮定する必要がある．証明をするには $X(\omega)$ の P 可測性から $X_T(\omega)$ の P 可測をいえばよいが，これは容易である．

$X(\omega)$ を S 値確率変数，$Y(\omega)$ を T 値確率変数とする．$Y(\omega)$ の値が $X(\omega)$ の値で定まるとき，すなわち

$$X(\omega_1) = X(\omega_2) \Longrightarrow Y(\omega_1) = Y(\omega_2)$$

のとき，$Y(\omega)$ は $X(\omega)$ の**娘**，$X(\omega)$ は $Y(\omega)$ の**母**という．このときには $f: S \to T$ が存在して，

$$f(X(\omega)) = Y(\omega) \quad \text{すなわち} \quad f \circ X = Y$$

となる．実際，$s \in X(\Omega)$ に対しては，上の条件で $Y(X^{-1}(s))$ は T の中の1点集合となるから，これを $\{f(s)\}$ とかき，$s \in S - X(\Omega)$ に対しては $f(s) \in T$ を任意に定めたらよい．

$Y(\omega)$ は P 可分であるから，T の上の可算分離族 $\mathcal{A} \subset \mathcal{D}(PY^{-1})$ がある．

$$PY^{-1} = (PX^{-1})f^{-1} = P^X f^{-1}$$

により，$\mathcal{A} \subset \mathcal{D}(P^X f^{-1})$．これは $f(s)$ が S の上の P^X 可分関数であること，いいかえれば，$f(s)$ が確率空間 (S, P^X) (これは可分完全) の上の T 値確率変数であることを意味する．

もし $Y(\omega)$ が T 値位相確率変数であれば，

$$\mathcal{B}(T) \subset \mathcal{D}(PY^{-1}) = \mathcal{D}(P^X f^{-1})$$

が得られ，$f(s)$ が (S, P^X) の上の T 値位相確率変数となる．

また上の $PY^{-1} = P^X f^{-1}$ は

$$\begin{aligned} \{\omega \mid Y(\omega) \in B\} \in \mathcal{D}(P) &\Longleftrightarrow \{s \mid f(s) \in B\} \in \mathcal{D}(P^X) \\ &\Longrightarrow P\{Y(\omega) \in B\} = P^X\{f(s) \in B\} \end{aligned}$$

を示している．

さて可分確率空間 (S, P^X) を X の**確率空間**といい，$f(s)$ を $Y(\omega) = f(X(\omega))$ の (S, P^X) の上の**表現**という．このような $f(s)$ は一意ではないが，$X(\Omega)$ の上では一意に定まり，しかも

$$P^X(X(\Omega)) = P(X^{-1}(X(\Omega))) = P(\Omega) = 1$$

により，(S, P^X) の上でほとんど確実に一意である．また $f(s)$ が (S, P^X) の上の任意の T 値 (位相) 確率変数ならば，$Y(\omega) = f(X(\omega))$ は (Ω, P) の上の T 値 (位相) 確率変数となり，$X(\omega)$ の娘となる．このようにして，$X(\omega)$ の娘だけを考察するときには，その表現をとることにより，(S, P^X) を基礎の確率空間として議論をすすめることができる．

表現が完全に一意でないことが気になるならば，値域の変更により $X(\Omega) = S$ の場合に帰着して議論をすすめたらよい．

$I: \Omega \to \Omega$ を恒等写像とすれば，$I(\omega)$ は (Ω, P) の上の Ω 値確率変数である．任意の確率変数 $X(\omega)$ は

$$X(\omega) = X(I(\omega))$$

とかけるから，$I(\omega)$ はすべての確率変数の共通の母である．$I(\omega) \equiv \omega$ であるから，ω 自身も確率変数と考えてよい．この意味で ω を**確率母変数**という．

可算個 (有限または可算無限個) の確率変数の結合変数を定義しよう．$X_n(\omega)$ $(n = 1, 2, \cdots)$ をそれぞれ S_n 値確率変数とする．

$$S=\prod_n S_n, \qquad p_k\colon S\longrightarrow S_k \text{ は射影写像} \quad (k=1,2,\cdots)$$

とする．$X_n(\omega)$ は P 可分であるから，S_n の上の可算分離族 $\mathscr{A}_n=\{A_{n1}, A_{n2}, \cdots\}$ があって

$$X_n^{-1}(A_{nk})\subset \mathscr{D}(P), \qquad k=1,2,\cdots.$$

さて (Ω, P) の上の S 値関数

$$X(\omega)=(X_1(\omega), X_2(\omega), \cdots)$$

も P 可分であるから，これは S 値確率変数であって，$X_n(\omega)$ $(n=1,2,\cdots)$ の**結合変数**という．$X(\omega)$ の P 可分性は，

$$\mathscr{A}=\{p_n^{-1}(A_{nk})\,|\,n,k=1,2,\cdots\}$$

が S の上の可算分離族で，しかも

$$\begin{aligned}X^{-1}(p_n^{-1}(A_{nk}))&=(p_n\circ X)^{-1}(A_{nk})\\&=X_n^{-1}(A_{nk})\in\mathscr{D}(P)\end{aligned}$$

となることによる．なお定理 3.1* により，P^X は $P^X|_{\sigma[\mathscr{A}]}$ の Lebesgue 拡大である．

さて上の S_n が位相空間 ($\mathscr{B}(S_n)$ は可算族で生成される) で，$X_n(\omega)$ が S_n 値位相確率変数であるとする．$S=\prod_n S_n$ を位相積空間とする．このとき S 値関数

$$X(\omega)=(X_1(\omega), X_2(\omega), \cdots)$$

は必ずしも P 可測でないから，S 値位相確率変数とはいえない．しかしすべての S_n が可算開基をもつならば，S も可算開基をもち，$X(\omega)$ は P 可測となって，S 値位相確率変数と考えられる．実際 S_n の可算開基を $\mathscr{U}_n=\{U_{n1}, U_{n2}, \cdots\}$ とすれば，

$$\mathscr{U}=\left\{U=\bigcap_{k=1}^{n}p_k^{-1}(U_{ki_k})\,\middle|\, n=1,2,\cdots, i_k=1,2,\cdots\ (k=1,2,\cdots)\right\}$$

が S の可算開基となり，上の U に対して，

$$X^{-1}(U)=\bigcap_{k=1}^{n}X_k^{-1}(U_{ki_k})\in\mathscr{D}(P)$$

となるが，S の中の任意の開集合 G は $\mathscr{U}$ の元の和 (したがって可算和) となるから，

$$X^{-1}(G)\in\mathscr{D}(P)$$

が得られ，これは X の P 可測性を意味する．

$\boldsymbol{R}^1$ は可算開基をもつから，可算個の実数値確率変数の (位相確率変数としての) 結合変数は考えられ，それは**確率ベクトル**である．

$[-\infty, \infty]$ を $\bar{\boldsymbol{R}}$ であらわし，**拡張された実直線**という．$\bar{\boldsymbol{R}}$ は可算開基をもつから，$\bar{\boldsymbol{R}}$ 位相確率変数が考えられる．これを**拡張された実確率変数**という．$\boldsymbol{R}^1\subset\bar{\boldsymbol{R}}$ で，$\boldsymbol{R}^1$ の位相は $\bar{\boldsymbol{R}}$ の位相の $\boldsymbol{R}^1$ への相対化であるから，値域の変更により，実確率変数を拡張された実確率変数の特別のものと考えることができる．$X(\omega)$ を拡張された実確率変数とするとき，P 可測集合 A に対して，積分

$$\int_A X(\omega)P(d\omega)$$

が考えられる．A の上で $X(\omega)\geqq 0$ のときには，上の積分は，$X(\omega)$ の P 可測性により確定し，$[0,\infty]$ の中の値をとる．一般の場合には

$$\int_A X^+(\omega)P(d\omega), \qquad \int_A X^-(\omega)P(d\omega)$$
$$(X^\pm(\omega)=\max(\pm X(\omega), 0))$$

の一方が有限のときにのみ確定し，$[-\infty,\infty]$ の中の値をとる．両方が無限のときには，上の積分は定義できない．また両方が有限のとき，すなわち

$$\int_A |X(\omega)|P(d\omega)<\infty$$

のときには，$X(\omega)$ は A の上で**可積分**といい，

$$\int_A X(\omega)P(d\omega)\in\boldsymbol{R}^1$$

となる．可積分というよりは，むしろ積分有限といった方がよいが，本講では伝統にしたがって可積分ということにする．

とくに

$$\int_\Omega X(\omega)P(d\omega)$$

が確定するとき，これを EX であらわし，$X(\omega)$ の**期待値** (または**平均値**) という．また

$$\int_A X(\omega)P(d\omega)$$

を $E(X, A)$ であらわす．

$$E(X, A) = E(X \cdot 1_A)$$

であるから，$E(X, A)$ の性質は EX の性質に帰着される．

EX の性質を列挙しておこう．(便宜上 $(\pm\infty)\cdot 0 = 0$ とする．)

定理 3.5 ($EX=EY$ は一方が確定すれば，他方も確定し，両者は相等しいことを意味する．)

(i) $E(1_A) = P(A)$.

(ii) $X(\omega) = Y(\omega)$ a.s. $\Longrightarrow EX = EY$.

(iii) $E(aX) = aEX$ (a は有限の定数)．

(iv) $X(\omega) \geqq 0$ a.s. とする．

$$EX = 0 \Longrightarrow X(\omega) = 0 \text{ a.s.},$$

$$EX < \infty \Longrightarrow X(\omega) < \infty \text{ a.s.}$$

(v) (**単調極限定理**) '$X_1(\omega) \leqq X_2(\omega) \leqq \cdots$ a.s. かつ $EX_1 > -\infty$'，または '$X_1(\omega) \geqq X_2(\omega) \geqq \cdots$ a.s. かつ $EX_1 < \infty$' ならば，$E\left(\lim_{n\to\infty} X_n\right) = \lim_{n\to\infty} EX_n$.

(vi) (**Fatou の補題**)

$$E\left(\inf_n X_n\right) > -\infty \Longrightarrow E\left(\liminf_{n\to\infty} X_n\right) \leqq \liminf_{n\to\infty} EX_n,$$

$$E\left(\sup_n X_n\right) < \infty \Longrightarrow E\left(\limsup_{n\to\infty} X_n\right) \geqq \limsup_{n\to\infty} EX_n,$$

$$E\left(\sup_n |X_n|\right) < \infty$$

$$\Longrightarrow E\left(\liminf_{n\to\infty} X_n\right) \leqq \liminf_{n\to\infty} EX_n \leqq \limsup_{n\to\infty} EX_n$$

$$\leqq E\left(\limsup_{n\to\infty} X_n\right).$$

(vii) (**有界収束定理**)

$$E\left(\sup_n |X_n|\right) < \infty,\quad X(\omega) = \lim_{n\to\infty} X_n(\omega) \text{ a.s.} \Longrightarrow EX = \lim_{n\to\infty} EX_n.$$

(viii) (**無限和との交換定理**)

$$X_n(\omega) \geqq 0 \ (n=1, 2, \cdots) \Longrightarrow E\left(\sum_n X_n\right) = \sum_n EX_n,$$

$$E\left(\sum_n |X_n|\right) \left(= \sum_n E|X_n|\right) < \infty \Longrightarrow E\left(\sum_n X_n\right) = \sum_n EX_n.$$

(ix) (**期待値の変換公式**) $X(\omega)$ を S 値確率変数 (S は一般の集合) とし，$f: S \to \bar{\boldsymbol{R}}$ を P^X 可測 (すなわち $f(s)$ は (S, P^X) の上の拡張された実確率変数) とすると，$f(X(\omega))$ は (Ω, P) の上の拡張された実確率変数で

$$E(f(X)) = E^X(f) \quad \left(= \int_S f(s) P^X(ds)\right)$$

がなりたつ．

証明 (i) から (viii) までは積分に関する周知の事実のいいかえにすぎないから，(ix) だけを証明する．積分の定義により，$f \geqq 0$ の場合だけを考察すればよい．$f = 1_B$ ($B \in \mathcal{D}(P^X)$) のときには，

$$E(1_B(X)) = P(X^{-1}(B)) = P^X(B) = E^X(1_B)$$

により変換公式がなりたつ．このような f の正係数1次結合 g に対してもなりたち，このような g の増大列の極限 h に対してもなりたつから，この変換公式は一般になりたつ．▮

例題 3.2 (i) $a_n > 0$, $\sum_n a_n < \infty$ とする．Borel-Cantelli の補題を用いて X_n $(n=1, 2, \cdots)$ が実確率変数で $P(|X_n| > a_n) < a_n$ $(n=1, 2, \cdots)$ ならば，$\sum_n X_n(\omega)$ はほとんど確実に絶対収束することを示せ．

(ii) $\boldsymbol{R}^n$ $(n=1, 2, \cdots, \infty)$ の上の正則確率測度は可分完全であることを示せ．

［ヒント］ これらの空間はすべて完備可分距離空間である．

(iii) 可分距離空間 T の中の値をとる P 可測関数は，位相確率変数と見なされ，したがってその確率法則は正則，可分かつ完全であることを示せ．

［ヒント］ この空間は可算開基をもつから，その Borel 集合族は可算族で生成される．

§3.3 分割と σ 加法族

(Ω, P) を基礎の確率空間とし，ω を Ω の一般元とすることは今まで通りである．Ω の空でない P 可測部分集合の族 Δ があって

$$\sum_{\xi \in \Delta} \xi = \Omega \qquad \text{(これは当然“}\Delta\text{ の 2 元が互いに素である”ことを含む)}$$

であるとき，Δ を Ω の**分割**という．

$$\Delta = \{A, A^c\} \qquad (A \in \mathcal{D}(P),\ A, A^c \neq \emptyset)$$

は分割である．$\Delta = \{\Omega\}$ も分割の極端なものである．$X(\omega)$ を Ω の上の関数とし

たとき

$$\Delta = \{X^{-1}(x) \mid x \in X(\Omega)\}$$

も分割である．この分割を関数 $X(\omega)$ による**分割**という．とくに $X(\omega)$ が P-可分のときには，$X(\omega)$ は確率変数となるから，Δ_X は確率変数 $X(\omega)$ による分割である．これは最も重要な分割である．

可算族 $\mathscr{A}=\{A_1, A_2, \cdots\}\subset\mathscr{D}(P)$ に対し

$$\xi = A_1'\cap A_2'\cap\cdots \qquad (A_n'=A_n \text{ または } A_n^c)$$

の形の空でない集合の全体を $\Delta_{\mathscr{A}}$ とかくと，$\Delta_{\mathscr{A}}$ も分割となる．これを可算族 $\mathscr{A}$ による分割という．可算族による分割を**可分分割**という．

定理 3.6 (i) 確率変数による分割は可分である．

(ii) 可分分割は確率変数による分割であり，しかもこの確率変数として実確率変数をとることができる．

証明 (i) $X(\omega)$ を S 値確率変数とし，$\Delta=\Delta_X$ とする．確率変数は，定義により，可分であるから，S の上の可算分離族 $\{B_n\}$ を $A_n=X^{-1}(B_n)$ が P 可測となるようにとれる．このとき Δ_X は可算族 $\{A_1, A_2, \cdots\}$ による分割と一致する．

(ii) Δ を $\mathscr{A}=\{A_1,A_2,\cdots\}\subset\mathscr{D}(P)$ による分割とする．実確率変数

$$X(\omega)=\sum_{n=1}^{\infty}3^{-n}1_{A_n}(\omega)$$

による分割 Δ_X が Δ と一致することを証明すればよい．任意の $\xi\in\Delta=\Delta_{\mathscr{A}}$ をとる．$\omega\in\xi$ に対しては，すべての n に対し，$1_{A_n}(\omega)$ の値が一定であるから，$X(\omega)$ の値も一定である．これは Δ の任意の元が Δ_X のある元に含まれることを意味する．上の $X(\omega)$ の式から

$$1_{A_n}(\omega)=[3^nX(\omega)]-3[3^{n-1}X(\omega)], \qquad n=1,2,\cdots$$

がでるから，Δ_X の任意の元は Δ のある元に含まれる．これらの事実から，Δ_X も Δ も分割であることに注意して，$\Delta=\Delta_X$ が容易に証明される．▌

分割 Δ' の任意の元が分割 Δ のある元に含まれているとき，Δ' は Δ の**細分**であるといい，

$$\Delta'\succ\Delta \qquad (\text{または } \Delta\prec\Delta')$$

であらわす．このときには Δ の任意の元は分割 Δ' のいくつか (一般に無限個) の元の直和となっている．

$$\Delta\prec\Delta',\quad \Delta'\prec\Delta''\Longrightarrow\Delta\prec\Delta'',$$
$$\Delta\prec\Delta',\quad \Delta'\prec\Delta\Longrightarrow\Delta'=\Delta$$

であるから，'$\prec$' は分割全体の系における半順序関係である．しかし一般に全順序関係ではない．$\Delta\prec\Delta'$ のとき Δ' は Δ より**大**，Δ は Δ' より**小**という．この大小関係にしたがって，分割の系 $\{\Delta_\lambda\}$ の**上限** $\bigvee_\lambda \Delta_\lambda$，**下限** $\bigwedge_\lambda\Delta_\lambda$ が定義される．すなわち Δ_λ のどれよりも大きい分割の中で最小のものが存在すれば，それは上限であり，Δ_λ のどれよりも小さい分割の中で最大なものが存在すれば，それは下限である．

定理 3.7 (i) 可算個の分割の系 $\{\Delta_1,\Delta_2,\cdots\}$ の上限 $\bigvee_n\Delta_n$ は確定し，

$$\bigvee_n \Delta_n=\left\{\bigcap_n A_n\neq\emptyset \,\middle|\, A_n\in\Delta_n\ (n=1,2,\cdots)\right\}$$

で与えられる．

(ii) $\{\Delta_n\}$ が増大分割列ならば，

$$\bigvee_n\Delta_n=\left\{\bigcap_n A_n\neq\emptyset\,\middle|\,A_n\in\Delta_n\ (n=1,2,\cdots),\ A_1\supset A_2\supset\cdots\right\}.$$

(このとき $\bigvee_n\Delta_n$ を $\{\Delta_n\}$ の**極限**といい，$\lim_n\Delta_n$ であらわす.)

(iii) $\{\Delta_n\}$ が減少分割列ならば，下限 $\bigwedge_n\Delta_n$ は確定し，

$$\bigwedge_n\Delta_n=\left\{\bigcup_n A_n\,\middle|\,A_n\in\Delta_n\ (n=1,2,\cdots),\ A_1\subset A_2\subset\cdots\right\}.$$

(このとき $\bigwedge_n\Delta_n$ を $\{\Delta_n\}$ の**極限**といい，$\lim_n\Delta_n$ であらわす.)

証明 (i) 等式の右辺を Δ とおく．Δ が分割であることは容易にわかる．

$$\bigcap_n A_n\subset A_k\in\Delta_k,\qquad k=1,2,\cdots$$

により，$\Delta\succ\Delta_k$ $(k=1,2,\cdots)$．もし Δ' が分割で，$\Delta'\succ\Delta_k$ $(k=1,2,\cdots)$ ならば，任意の $A'\in\Delta'$ に対し，A' を含む $A_k\in\Delta_k$ をとると

$$A'\subset\bigcap_k A_k\in\Delta,$$

これは $\Delta'\succ\Delta$ を意味する．これで $\Delta=\bigvee_n\Delta_n$ が証明された．

(ii) 分割の定義により，(i) の $\{A_n\}$ が減少列となるから，(ii) は (i) からすぐでる．

(iii) 等式の右辺を Δ とおく．Δ の2元 $\xi=\bigcup_n A_n$，$\eta=\bigcup_n B_n$ が共通点をも

てば，

$$\bigcup_{m,n} A_m \cap B_n = \xi \cap \eta \neq \phi.$$

したがって $A_m \cap B_n \neq \phi$ となる m, n がある．m, n の大きい方を k_0 とすれば，$k > k_0$ に対し

$$A_k \cap B_k \supset A_m \cap B_n \neq \phi.$$

$A_k, B_k \in \varDelta_k$ により，$A_k = B_k$．$\{A_n\}$, $\{B_n\}$ は増大列であるから，

$$\xi = \bigcup_k A_k = \bigcup_{k>k_0} A_k, \qquad \eta = \bigcup_k B_k = \bigcup_{k>k_0} B_k.$$

上に証明した $A_k = B_k$ $(k > k_0)$ により，$\xi = \eta$．これで $\varDelta$ の元が互いに素であることが証明された．任意の $A_1 \in \varDelta_1$ に対して，A_1 を含む $A_2 \in \varDelta_2$，A_2 を含む $A_3 \in \varDelta_3$，というように

$$A_1 \subset A_2 \subset \cdots, \qquad A_n \in \varDelta_n \quad (n = 1, 2, \cdots)$$

をとることができる（仮定 $\varDelta_1 \succ \varDelta_2 \succ \cdots$ による）．したがって

$$A_1 \subset \bigcup_n A_n \in \varDelta.$$

すべての $A_1 \in \varDelta_1$ の和は Ω であるから，$\varDelta$ のすべての元の和も Ω に等しい．これで $\varDelta$ が分割であることが分った．$\varDelta = \bigwedge_n \varDelta_n$ は，(i) の証明に用いた論法で，容易に証明される．▌

定理 3.8 可算個の可分分割の系の上限は可分である．

証明 $\varDelta_n$ が可算族 $\mathscr{A}_n$ による分割とすると，$\bigvee_n \varDelta_n$ は

$$\mathscr{A} = \bigcup_n \mathscr{A}_n \qquad (\text{集合族としての和})$$

による分割であることに注意すればよい．▌

$\varDelta_n$ が確率変数 $X_n(\omega)$ による分割であれば，$\bigvee_n \varDelta_n$ は明らかに結合変数 $X(\omega) = (X_1(\omega), X_2(\omega), \cdots)$ による分割と一致することは定義から明らかである．この事実と定理 3.6 からも定理 3.8 がすぐにでる．

定理 3.8 の特別の場合として，可分分割の増大列の極限は可分であることがわかる．しかし可分分割の減少列の極限は必ずしも可分ではない．反例は次節にのべる（例題 3.4 (vi)）．

定理 3.9 $X(\omega)$, $Y(\omega)$ を確率変数とする．

$$Y(\omega) \text{ が } X(\omega) \text{ の娘である} \Longleftrightarrow \varDelta_Y \prec \varDelta_X.$$

証明 娘の定義，分割の大小の定義から容易に証明される．▌

分割 $\varDelta$ に対して，$\varDelta$ 値関数 $X_\varDelta$ を

$$X_\varDelta(\omega) = \omega \text{ を含む } \varDelta \text{ の元}$$

で定義する．明らかに

$$\varDelta = \varDelta_{X_\varDelta}$$

がなりたつ．したがって $\varDelta$ の可分性と $X_\varDelta$ の P 可分性（すなわち $X_\varDelta(\omega)$ が $\varDelta$ 値確率変数となること）は同値である．このとき $X_\varDelta(\omega)$ を **$\varDelta$ に対応する確率変数**という．

分割についてはこの辺にとどめて，つぎに σ 加法族についてのべる．ここで σ 加法族というのは P 可測集合からなる σ 加法族である．$\mathcal{D}(P)$ は最大の σ 加法族で，$\{\phi, \Omega\}$ は最小の σ 加法族である．また

$$\mathbf{2} = \{A \in \mathcal{D}(P) \mid P(A) = 0 \text{ または } 1\}$$

も σ 加法族である．

σ 加法族の全体は包含関係に関して半順序系をつくる．$\mathcal{B} \subset \mathcal{B}'$ のとき $\mathcal{B}'$ は $\mathcal{B}$ より大きい，$\mathcal{B}$ は $\mathcal{B}'$ より小さいという．σ 加法族の系 $\mathcal{B}_\lambda$ $(\lambda \in \Lambda)$ に対して，すべての $\mathcal{B}_\lambda$ より大きい σ 加法族のうち最小のものを

$$\bigvee_\lambda \mathcal{B}_\lambda \quad \text{または} \quad \sup_\lambda \mathcal{B}_\lambda$$

であらわし，すべての $\mathcal{B}_\lambda$ より小さい σ 加法族のうち最大のものを

$$\bigwedge_\lambda \mathcal{B}_\lambda \quad \text{または} \quad \inf_\lambda \mathcal{B}_\lambda$$

であらわす．明らかに

$$\bigvee_\lambda \mathcal{B}_\lambda = \sigma\Big[\bigcup_\lambda \mathcal{B}_\lambda\Big], \qquad \bigwedge_\lambda \mathcal{B}_\lambda = \bigcap_\lambda \mathcal{B}_\lambda.$$

ここで $\bigcup_\lambda \mathcal{B}_\lambda$, $\bigcap_\lambda \mathcal{B}_\lambda$ はそれぞれ集合族としての和，交わりをあらわす．

σ 加法族の列 $\{\mathcal{B}_n\}$ に対して，その上極限，下極限を

$$\limsup_n \mathcal{B}_n = \bigwedge_n \bigvee_{k>n} \mathcal{B}_k, \qquad \liminf_n \mathcal{B}_n = \bigvee_n \bigwedge_{k>n} \mathcal{B}_k$$

で定義する．$\{\mathcal{B}_n\}$ が増大または減少のときには，この両極限は一致し，$\lim_n \mathcal{B}_n$ であらわす．$\{\mathcal{B}_n\}$ が増大か減少かに応じて

$$\lim_n \mathcal{B}_n = \bigvee_n \mathcal{B}_n \text{ または } \bigwedge_n \mathcal{B}_n$$

となる．

分割 Δ に対して，Δ のいくつか(無限個でもよい)の元の和であらわされるような P 可測集合の全体を Δ で定まる σ 加法族といい，$\mathcal{B}_\Delta$ であらわす．

$\mathcal{B}$ を σ 加法族，Δ を分割とする．

$$\Delta \subset \mathcal{B} \subset \mathcal{B}_\Delta$$

のとき，$\mathcal{B}$ は分割 Δ に**従属**するという．分割 Δ に従属する σ 加法族の全体を $\boldsymbol{B}(\Delta)$ であらわす．明らかに $\mathcal{B}_\Delta$ は Δ に従属する．

$X(\omega)$ が確率変数のとき，

$$X^{-1}(\mathcal{D}(P^X)) = \{X^{-1}(E) \mid E \in \mathcal{D}(P^X)\}$$

は σ 加法族である．これを $\bar{\sigma}[X]$ であらわす．明らかに

$$\Delta = \Delta_X \Longrightarrow \mathcal{B}_\Delta = \bar{\sigma}[X]$$

がなりたつ．したがって $\bar{\sigma}[X]$ は Δ_X に従属する．

$X(\omega)$ が位相確率変数(値域 S)のときにも，$\bar{\sigma}[X]$ は上のように定義することはいうまでもないが，このほかに $\sigma[X]$ を

$$\sigma[X] = X^{-1}(\mathcal{B}(S)) = \{X^{-1}(E) \mid E \in \mathcal{B}(S)\}$$

で定義する．明らかに

$$\Delta_X \subset \sigma[X] \subset \bar{\sigma}[X] = \mathcal{B}_{\Delta_X}$$

であるから，$\sigma[X]$ も $\bar{\sigma}[X]$ も Δ_X に従属する．

σ 加法族 $\mathcal{B}_1, \mathcal{B}_2$ が**同等**であるとは

$$\mathcal{B}_1 \vee \boldsymbol{2} = \mathcal{B}_2 \vee \boldsymbol{2}$$

となることである．これは $\mathcal{B}_1, \mathcal{B}_2$ の一方の任意の元が他方のある元と P 零集合だけしか違わないことを意味する．同等性は同値関係である．これより強い同値関係として**ほとんど一致する**という関係がある．$\mathcal{B}_1, \mathcal{B}_2$ がほとんど一致するとは，適当な P 零集合 N をとれば，

$$\mathcal{B}_1 \cap (\Omega - N) = \mathcal{B}_2 \cap (\Omega - N)$$

となることである．ここで $\mathcal{B} \cap A$ は $\{B \cap A \mid B \in \mathcal{B}\}$ をあらわす．

例題 3.3 (i) $f: S \to T$，$\mathcal{B}$ は T の上の σ 加法族とするとき，$f^{-1}(\mathcal{B})$ は S の上の σ 加法族であることを示せ．

(ii) $f: S \to T$，$\mathcal{B}$ は S の上の σ 加法族とする．

$$f[\mathcal{B}] = \{B \subset T \mid f^{-1}(B) \in \mathcal{B}\}$$

は T の上の σ 加法族であることを示せ．また

$$f(\mathcal{B}) = \{f(B) \mid B \in \mathcal{B}\}$$

は T の上の σ 加法族とは限らないことを示す反例をあげよ．

(iii) S を位相空間の系 $S_\lambda\ (\lambda \in \Lambda)$ の位相積空間とし，

$$\mathcal{B} = \bigvee_\lambda p_\lambda^{-1}(\mathcal{B}(S_\lambda)) \qquad (p_\lambda : S \to S_\lambda \text{ は射影})$$

とおく．一般に

$$\mathcal{B}(S) \supset \mathcal{B}$$

であり，とくに Λ が可算集合，すべての S_λ が可算開基をもてば，$\mathcal{B}(S) = \mathcal{B}$ となることを証明せよ．

(iv) 前問ですべての S_λ が少なくとも2点を含み，Λ が非可算のときには，$\mathcal{B}(S) \neq \mathcal{B}$ となることを示せ．

[ヒント] $B \subset S$ が**可算決定**であるとは，適当な可算集合 $\Lambda_0 \subset \Lambda$ (B に関係してよい)が存在して，$s, s' \in S$ に対し，

$$p_\lambda(s) = p_\lambda(s') \quad (\lambda \in \Lambda_0) \Longrightarrow 1_B(s) = 1_B(s')$$

がなりたつことである．可算決定の集合 $B \subset S$ の全体 $\mathcal{B}_0$ は S の上の σ 加法族をつくり，

$$\mathcal{B}_0 \supset p_\lambda^{-1}(\mathcal{B}(S_\lambda)),\quad \lambda \in \Lambda \quad したがって \quad \mathcal{B}_0 \supset \mathcal{B}.$$

これは $\mathcal{B}$ の元はすべて可算決定であることを意味する．さて S の1点集合は閉集合で $\mathcal{B}(S)$ に属するが，すべての S_λ が2点以上を含むから，可算決定でなく，$\mathcal{B}$ には属さない．

(v) 次のことを証明せよ．

$$\mathcal{B}_n \in \boldsymbol{B}(\Delta_n) \quad (n=1, 2, \cdots) \Longrightarrow \bigvee_n \mathcal{B}_n \in \boldsymbol{B}\Bigl(\bigvee_n \Delta_n\Bigr).$$

(vi) $X(\omega)$ が位相確率変数のとき，$\sigma[X]$ と $\bar{\sigma}[X]$ とは同等であることを証明せよ．

(vii) すべての $\lambda \in \Lambda$ に対し $\mathcal{B}_\lambda$ と $\mathcal{B}_\lambda'$ とが同等ならば，$\bigvee_\lambda \mathcal{B}_\lambda$ と $\bigvee_\lambda \mathcal{B}_\lambda'$ とが同等であることを示せ．

[ヒント] $\Bigl(\bigvee_\lambda \mathcal{B}_\lambda\Bigr) \vee \boldsymbol{2} = \bigvee_\lambda (\mathcal{B}_\lambda \vee \boldsymbol{2})$ に注意せよ．

(viii) $X(\omega)$ が $X_1(\omega), X_2(\omega), \cdots$ の結合変数であるとき，P^X はその $\mathcal{B} \equiv \bigvee_n p_n^{-1}(\mathcal{D}(P^{X_n}))$ への制限の Lebesgue 拡大であることを示せ．

[ヒント] 結合変数の定義の際に用いた $\mathscr{A}$ について $\sigma[\mathscr{A}]\subset\bigvee_n p_n^{-1}(\mathscr{D}(P^{X_n}))\subset\mathscr{D}(P^X)$ に注意し，定理3.1* を用いよ．

§3.4 独立

前節と同様に P 可測集合からなる σ 加法族を簡単に σ 加法族という．σ 加法族の系 $\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ が**独立**であるとは，任意の有限集合 $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}\subset\Lambda$ に対し

$$P\left(\bigcap_{i=1}^{n} B_i\right)=\prod_{i=1}^{n} P(B_i),\qquad B_i\in\mathscr{B}_{\lambda_i}\quad(i=1, 2, \cdots, n)$$

がなりたつことである．

定理 3.10 (i) $\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ の任意の有限部分系が独立ならば，$\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ が独立である．

(ii) 独立な σ 加法族の系の部分系は独立である．

(iii) すべての $\lambda\in\Lambda$ に対し $\mathscr{B}_\lambda$ と $\mathscr{B}_\lambda'$ とが同等ならば，$\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ の独立性と $\{\mathscr{B}_\lambda', \lambda\in\Lambda\}$ の独立性とは同値である．

(iv) すべての $\lambda\in\Lambda$ に対し $\mathscr{B}_\lambda'\subset\mathscr{B}_\lambda$ ならば，$\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ の独立性から $\{\mathscr{B}_\lambda', \lambda\in\Lambda\}$ の独立性がでる．

証明 定義から明らか．∎

定理 3.11 $\mathscr{A}_\lambda$ $(\lambda\in\Lambda)$ はすべて乗法族とし，$\mathscr{B}_\lambda=\sigma[\mathscr{A}_\lambda]$ とする．任意の集合 $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}\subset\Lambda$ に対し

$$P\left(\bigcap_{i=1}^{n} A_i\right)=\prod_{i=1}^{n} P(A_i),\qquad A_i\in\mathscr{A}_{\lambda_i}\quad(i=1, 2, \cdots, n)$$

がなりたつならば，$\mathscr{B}_\lambda$ $(\lambda\in\Lambda)$ は独立である．

証明 $m=0, 1, 2, \cdots$ に関係するつぎの命題を帰納法で証明する．

(I_m) "任意の n $(\geqq m)$ 個の異なる元 $\lambda_1, \lambda_2, \cdots, \lambda_n\in\Lambda$ に対し

$$P\left(\bigcap_{i=1}^{m} B_i\cap\bigcap_{j=m+1}^{n} A_j\right)=\prod_{i=1}^{m} P(B_i)\prod_{j=m+1}^{n} P(A_j),\qquad B_i\in\mathscr{B}_{\lambda_i},\ \ A_j\in\mathscr{A}_{\lambda_j}$$

がなりたつ．($m=0, n$ のときには，それぞれ B_i に関する部分，A_j に関する部分は消える．)"

もしこれが証明できれば，$n=m$ とおくと，定理の結論が得られる．

(I_0) は定理の仮定そのものである．(I_m) が成立すると仮定して，(I_{m+1}) を証明しよう．任意の n $(\geqq m+1)$ 個の異なる元 $\lambda_1, \lambda_2, \cdots, \lambda_n\in\Lambda$ に対し，

$$P\left(\bigcap_{i=1}^{m+1} B_i\cap\bigcap_{j=m+2}^{n} A_j\right)=\prod_{i=1}^{m+1} P(B_i)\prod_{j=m+2}^{n} P(A_j),\qquad B_i\in\mathscr{B}_{\lambda_i},\ \ A_j\in\mathscr{A}_{\lambda_j}$$

がなりたつことを (I_m) から導けばよい．

$B_2, B_3, \cdots, B_{m+1}, A_{m+2}, \cdots, A_n$ を任意に固定し，上の式がなりたつような B_1 の全体 $\mathscr{B}$ が $\mathscr{B}_{\lambda_1}$ を含むことをいえば，定理の証明は完了する．確率測度 P の σ 加法性により，$\mathscr{B}$ は固有差，可算直和に対して閉じている．また帰納法の仮定により，$B_1=\Omega$ または $B_1\in\mathscr{A}_{\lambda_1}$ のときには上の式がなりたつから，$\mathscr{B}$ は $\mathscr{A}_{\lambda_1}$ を含む Dynkin 族である．$\mathscr{A}_{\lambda_1}$ は乗法族であるから，Dynkin 族定理により，$\mathscr{B}$ は $\sigma[\mathscr{A}_{\lambda_1}]$ すなわち $\mathscr{B}_{\lambda_1}$ を含む．∎

定理 3.12

$$\Lambda=\sum_{\mu\in M}\Lambda_\mu$$

とし，$\lambda\in\Lambda$ に対して σ 加法族 $\mathscr{B}_\lambda$ が与えられているとし，

$$\mathscr{B}_\mu'=\bigvee_{\lambda\in\Lambda_\mu}\mathscr{B}_\lambda,\qquad \mu\in M$$

とおく．このとき

$$\mathscr{B}_\lambda\ (\lambda\in\Lambda)\text{ が独立}\iff\begin{cases}\text{すべての }\mu\in M\text{ に対し }\mathscr{B}_\lambda\ (\lambda\in\Lambda_\mu)\text{ が独立,}\\ \mathscr{B}_\mu'\ (\mu\in M)\text{ が独立.}\end{cases}$$

証明 Λ が有限集合(したがって Λ_μ も M も有限)の場合を考えれば十分である．$\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ が独立であるとせよ．任意の $\mu\in M$ に対し，$\{\mathscr{B}_\lambda, \lambda\in\Lambda_\mu\}$ は $\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ の部分系であるから独立である．

$$A_\mu'=\bigcap_{\lambda\in\Lambda_\mu}B_\lambda,\qquad B_\lambda\in\mathscr{B}_\lambda$$

の形の集合 A_μ' の全体 $\mathscr{A}_\mu'$ は乗法族で，$\mathscr{B}_\mu'$ は $\mathscr{A}_\mu'$ で生成される．しかも $\{\mathscr{B}_\lambda, \lambda\in\Lambda\}$ の独立性により，

$$P\left(\bigcap_{\mu\in M}A_\mu'\right)=P\left(\bigcap_{\mu\in M}\bigcap_{\lambda\in\Lambda_\mu}B_\lambda\right)=\prod_{\mu\in M}\prod_{\lambda\in\Lambda_\mu}P(B_\lambda)=\prod_{\mu\in M}P(A_\mu')$$

となり，前定理により，$\{\mathscr{B}_\mu', \mu\in M\}$ が独立となる．これで '$\Longrightarrow$' の部分が証明された．'$\Longleftarrow$' を証明するために，$B_\lambda\in\mathscr{B}_\lambda$ $(\lambda\in\Lambda)$ を任意にとり，これに対し，$A_\mu'\in\mathscr{B}_\mu'$ を上のように定義すると，仮定により

$$P(A_\mu')=\prod_{\lambda\in\Lambda_\mu}P(B_\lambda),\qquad P\left(\bigcap_{\mu\in M}A_\mu'\right)=\prod_{\mu\in M}P(A_\mu')$$

となる．したがって

$$P\Big(\bigcap_{\lambda\in\Lambda}B_\lambda\Big)=P\Big(\bigcap_{\mu\in M}A_\mu'\Big)=\prod_{\mu\in M}\prod_{\lambda\in\Lambda_\mu}P(B_\lambda)=\prod_{\lambda\in\Lambda}P(B_\lambda)$$

となり，'$\Longleftarrow$'が証明された．▮

定理 3.13 (Kolmogorov の 0-1 法則)　$\mathcal{B}_1, \mathcal{B}_2, \cdots$ が独立ならば，

$$\limsup_{n\to\infty}\mathcal{B}_n\Big(=\bigwedge_k\bigvee_{n>k}\mathcal{B}_n\Big)\subset \boldsymbol{2},$$

すなわち左辺に属する任意の集合 B に対し

$$P(B)=0 \text{ または } 1.$$

証明　$\mathcal{C}_k=\bigvee_{n>k}\mathcal{B}_n,\quad \mathcal{C}=\bigwedge_k\mathcal{C}_k=\bigwedge_k\bigvee_{n>k}\mathcal{B}_n,\quad \mathcal{B}=\bigvee_n\mathcal{B}_n$

とおく．前定理で

$$\Lambda=\{1,2,3,\cdots\},\qquad M=\{1,2,\cdots,k+1\},$$

$$\Lambda_1=\{1\},\quad \Lambda_2=\{2\},\quad \cdots,\quad \Lambda_k=\{k\},\quad \Lambda_{k+1}=\{k+1,k+2,\cdots\}$$

とおいて見ると，$\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_k, \mathcal{C}_k$ が独立であることがわかる．$\mathcal{C}\subset\mathcal{C}_k$ により $\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_k, \mathcal{C}$ が独立である．これはすべての k に対してなりたつから，定理3.10 (i) により，$\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{C}$ も独立である．再び前定理を用いて，$\mathcal{B}, \mathcal{C}$ が独立．$\mathcal{C}\subset\mathcal{B}$ により，$\mathcal{C}$ とそれ自身とが独立．したがって，$B\in\mathcal{C}$ に対し

$$P(B)=P(B\cap B)=P(B)P(B).$$

これから $P(B)=0$ または 1．▮

$\{X_\lambda(\omega), \lambda\in\Lambda\}$ を確率変数の系とする．σ 加法族の系 $\{\bar\sigma[X_\lambda], \lambda\in\Lambda\}$ が独立ならば，$\{X_\lambda(\omega), \lambda\in\Lambda\}$ は**独立**であるという．

定理 3.14　(i)　$\{X_\lambda(\omega), \lambda\in\Lambda\}$ が独立，おのおのの $\lambda\in\Lambda$ に対し $Y_\lambda(\omega)$ が $X_\lambda(\omega)$ の娘ならば，$\{Y_\lambda(\omega), \lambda\in\Lambda\}$ も独立である．

(ii)　"$\{X_{\lambda n}(\omega), \lambda\in\Lambda, n=1,2,\cdots\}$ が独立である" ことと，"おのおのの λ に対し $\{X_{\lambda1}(\omega), X_{\lambda2}(\omega), \cdots\}$ が独立，かつ結合変数の系

$$X_\lambda(\omega)=(X_{\lambda1}(\omega), X_{\lambda2}(\omega), \cdots),\qquad \lambda\in\Lambda$$

が独立である" こととは同等である．

(iii)　$\{X_1(\omega), X_2(\omega), \cdots\}$ が独立，その結合変数を $X(\omega)$ とすれば，P^X は P^{X_n} $(n=1,2,\cdots)$ の完備直積測度である．

証明　(i)　$Y(\omega)$ が $X(\omega)$ の娘ならば，$\bar\sigma[Y]\subset\bar\sigma[X]$ となることに注意すれば，定理3.10 (iv) により明らか．

(ii)　$X(\omega)=(X_1(\omega), X_2(\omega), \cdots)$ のとき，$\bar\sigma[X]$ が $\bigvee_n\bar\sigma[X_n]$ と同等なことをいえば，(ii) は定理3.12の特別の場合である．例題3.3 (viii) により $\mathcal{D}(P^X)$ は $\bigvee_n p_n^{-1}(\mathcal{D}(P^{X_n}))$ と (P^X に関し) 同等である．ゆえに $\bar\sigma[X]=X^{-1}(\mathcal{D}(P^X))$ は

$$X^{-1}\Big(\bigvee_n p_n^{-1}(\mathcal{D}(P^{X_n}))\Big)=\bigvee_n X_n^{-1}(\mathcal{D}(P^{X_n}))=\bigvee_n\bar\sigma[X_n]$$

と (P に関して) 同等である．

(iii)　$B_n\in\mathcal{D}(P^{X_n})$ $(n=1,2,\cdots)$ に対し

$$\begin{aligned}P^X(B_1\times B_2\times\cdots)&=P(X^{-1}(B_1\times B_2\times\cdots))\\&=P\Big(\bigcap_n X_n^{-1}(B_n)\Big)=\prod_n P(X_n^{-1}(B_n))=\prod_n P^{X_n}(B_n).\end{aligned}$$

これは P^X の $\bigvee_n p_n^{-1}(\mathcal{D}(P^{X_n}))$ への制限 Q が直積測度 $\prod_n P^{X_n}$ であることを意味する．例題3.3 (viii) により，P^X は Q の Lebesgue 拡大であるから，P^X は P^{X_n} $(n=1,2,\cdots)$ の完備直積である．▮

$\{X_\lambda(\omega), \lambda\in\Lambda\}$ が位相確率変数の系であれば，その独立性は $\sigma[X_\lambda]$ $(\lambda\in\Lambda)$ の独立性によって定義してもよいことは，$\sigma[X_\lambda]$ と $\bar\sigma[X_\lambda]$ とが同等であることと，定理3.10 (iii) とから直ちにわかる．また定理3.11により，任意の $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}\subset\Lambda$ に対し

$$P\Big(\bigcap_{i=1}^n X_{\lambda_i}^{-1}(G_i)\Big)=\prod_{i=1}^n P(X_{\lambda_i}^{-1}(G_i))\qquad (G_i \text{ は } X_{\lambda_i} \text{ の値域の開部分集合})$$

がなりたてば，$\{X_\lambda(\omega), \lambda\in\Lambda\}$ は独立となる．

P 可測集合の系 $\{A_\lambda, \lambda\in\Lambda\}$ (事象系と考えてもよい) が**独立**であるとは，確率変数系 $\{1_{A_\lambda}(\omega), \lambda\in\Lambda\}$ が独立であることと定義する．これはまた σ 加法族の系

$$\mathcal{B}_\lambda=\{\phi, A_\lambda, A_\lambda^c, \Omega\},\qquad \lambda\in\Lambda$$

が独立であることといっても同じである．定理3.11によれば，$\{A_1, A_2, \cdots, A_n\}$ が独立であるためには，任意の $k=2,3,\cdots,n$ と任意の $1\leqq i_1<i_2<\cdots<i_k\leqq n$ に対し

$$P\Big(\bigcap_{\nu=1}^k A_{i_\nu}\Big)=\prod_{\nu=1}^k P(A_{i_\nu})$$

となることが十分である．(もちろん必要でもある．)　この条件は定理1.21の条

件 (iii) にほかならない．

有限試行の場合の平均値の乗法性 (定理 1.23) は一般の場合にもつぎの形でなりたつ．

定理 3.15 (平均値の乗法性)　$X_1(\omega), X_2(\omega), \cdots, X_n(\omega)$ が独立な実確率変数 (複素確率変数でもよい) であるとする．

(i)　$X_k(\omega) \geqq 0$ $(k=1,2,\cdots,n)$ ならば，$E(\prod_k X_k) = \prod_k EX_k$.

(ii)　EX_k $(k=1,2,\cdots,n)$ がすべて有限確定ならば，同じ等式がなりたつ．

証明　$X_k(\omega)$ が P 可測集合 A_k の指示関数のときには，(i) の等式は

$$EX_k = P(A_k), \qquad E\left(\prod_k X_k\right) = P\left(\bigcap_k A_k\right)$$

から明らかである．これから，$X_k(\omega)$ がこのような指示関数の正係数の1次結合のときにもなりたつことが平均値の加法性からわかり，さらに平均値に関する単調極限定理を用いて (i) の一般の場合が証明できる．(ii) を (i) から導くには，$X_k(\omega)$ が実確率変数のときには分解

$$X_k = X_k^+ - X_k^-,$$

さらに複素確率変数のときには分解

$$X_k = \mathrm{Re}\,(X_k) + i\,\mathrm{Im}\,(X_k)$$

を念頭において，平均値の加法性を用いる．▮

確率変数の独立性は極限にまで遺伝する．正確にいうと，つぎの定理がなりたつ．

定理 3.16　おのおのの $\lambda \in \Lambda$ に対し，$X_\lambda^{(n)}(\omega), X_\lambda(\omega)$ が距離空間 S_λ の中の値をとる位相確率変数とする．もしおのおのの n に対し $\{X_\lambda^{(n)}(\omega), \lambda \in \Lambda\}$ が独立であり，おのおのの $\lambda \in \Lambda$ に対し，$\{X_\lambda^{(n)}(\omega), n=1,2,\cdots\}$ が $X_\lambda(\omega)$ にほとんど確実に収束するならば，$\{X_\lambda(\omega), \lambda \in \Lambda\}$ も独立である．

証明　定理 3.10 (i) により，Λ が有限集合の場合に証明すればよい．$f_\lambda : S_\lambda \to \boldsymbol{R}^1$ を有界連続とする．定理 3.14 (i) により，独立性は娘に遺伝するから，おのおのの n に対し，

$$Y_\lambda^{(n)}(\omega) = f_\lambda(X_\lambda^{(n)}(\omega)), \qquad \lambda \in \Lambda$$

も独立である．したがって平均値の乗法性により

$$E\left(\prod_\lambda Y_\lambda^{(n)}\right) = \prod_\lambda E(Y_\lambda^{(n)}).$$

$n \to \infty$ として，有界収束定理により

$$E\left(\prod_\lambda f_\lambda(X_\lambda(\omega))\right) = \prod_\lambda E(f_\lambda(X_\lambda(\omega))).$$

距離空間では，任意の開集合 G の指示関数 1_G は 0 と 1 との間の値をとる連続関数の列の極限としてあらわされるから，再び有界収束定理により上の等式が $f_\lambda = 1_{G_\lambda}$ (G_λ は S_λ の開部分集合) のときにもなりたつことがわかる．これは

$$P\left(\prod_\lambda X_\lambda^{-1}(G_\lambda)\right) = \prod_\lambda P(X_\lambda^{-1}(G_\lambda))$$

を意味する．したがって $\{X_\lambda(\omega), \lambda \in \Lambda\}$ は独立である．▮

分割の系 $\{\varDelta_\lambda, \lambda \in \Lambda\}$ が独立であるとは，それぞれの分割で定まる σ 加法族の系 $\{\mathcal{B}_{\varDelta_\lambda}, \lambda \in \Lambda\}$ が独立であることと定義する．明らかに

$$\{\varDelta_\lambda, \lambda \in \Lambda\} \text{ が独立} \iff \{X_{\varDelta_\lambda}, \lambda \in \Lambda\} \text{ が独立},$$
$$\{X_\lambda, \lambda \in \Lambda\} \text{ が独立} \iff \{\varDelta_{X_\lambda}, \lambda \in \Lambda\} \text{ が独立}.$$

事象系 $\{\alpha_\lambda(\omega), \lambda \in \Lambda\}$ の独立性は対応する集合の系 $\{\{\alpha_\lambda\}, \lambda \in \Lambda\}$ の独立性によって定義される．独立な事象列については Borel-Cantelli の補題はつぎのように精密化される．

定理 3.17 (Borel の定理)　$\alpha_1, \alpha_2, \cdots$ が独立ならば，

$$\sum_n P\{\alpha_n\} < \infty \implies P\{\alpha_n \ \text{i.o.}\} = 0, \quad P\{\alpha_n^{\urcorner} \ \text{f.e.}\} = 1,$$
$$\sum_n P\{\alpha_n\} = \infty \implies P\{\alpha_n \ \text{i.o.}\} = 1, \quad P\{\alpha_n^{\urcorner} \ \text{f.e.}\} = 0.$$

証明　前半は定理 3.4 (Borel-Cantelli の定理) に含まれている．後半を示そう．$A_n = \{\alpha_n\}$ とおいて

$$P\{\alpha_n^{\urcorner} \ \text{f.e.}\} = P\left\{\bigcup_k \bigcap_{n \geqq k} A_n^c\right\} \leqq \sum_k P\left(\bigcap_{n \geqq k} A_n^c\right),$$

$$P\left(\bigcap_{n \geqq k} A_n^c\right) \leqq P\left(\bigcap_{n=k}^m A_n^c\right) = \prod_{n=k}^m P(A_n^c) \qquad (\text{独立性})$$

$$= \prod_{n=k}^m (1 - P(A_n)).$$

$\sum P(A_n) = \infty$ により，上の最後の積は，$m \to \infty$ のとき，0 に収束し，$P\{\alpha_n^{\urcorner} \ \text{f.e.}\}$

$=0$ を得る．したがって $P\{\alpha_n \text{ i.o.}\}=1$. ∎

例題 3.4　(i)　すべての k に対し，$\mathcal{B}_k$ が $\bigvee_{n>k}\mathcal{B}_n$ と独立（一般に $\{\mathcal{B},\mathcal{B}'\}$ が独立のとき，$\mathcal{B}$ は $\mathcal{B}'$ と独立という）ならば，$\mathcal{B}_1,\mathcal{B}_2,\cdots$ が独立となることを示せ．

［ヒント］　$\{\mathcal{B}_1, \mathcal{C}_1\equiv\bigvee_{n>1}\mathcal{B}_n\}$ の独立性と，$\{\mathcal{B}_2,\mathcal{C}_2\equiv\bigvee_{n>2}\mathcal{B}_n\}$ の独立性とから，$\mathcal{C}_1=\mathcal{B}_2\vee\mathcal{C}_2$ に注意して定理 3.12 を用いると，$\{\mathcal{B}_1,\mathcal{B}_2,\mathcal{C}_2\}$ の独立性がわかる．この論法をくりかえして，任意の k に対して $\{\mathcal{B}_1,\mathcal{B}_2,\cdots,\mathcal{B}_k,\mathcal{C}_k\}$ の独立性が得られ，したがってもちろん $\{\mathcal{B}_1,\mathcal{B}_2,\cdots,\mathcal{B}_k\}$ が独立となる．

(ii)　$X_1,X_2,\cdots,X_n$ が独立な実確率変数のとき，

$$Y_1 = X_1+X_2+\cdots+X_k,\qquad Y_2=X_{k+1}+X_{k+2}+\cdots+X_n$$

が独立であることを証明せよ．

［ヒント］　$\sigma[Y_1]\subset\sigma[(X_1,\cdots,X_k)]=\bigvee_{i=1}^{n}\sigma[X_i]$，同様に $\sigma[Y_2]\subset\bigvee_{j=k+1}^{n}\sigma[X_j]$ に注意し，定理 3.12 を用いよ．

(iii)　実確率変数 $X(\omega)$ が可測 $\mathbf{2}$ ならば，$X(\omega)$ はほとんど確実にある定数 a に等しいことを証明せよ．

［ヒント］　$X^{-1}(-\infty,x]\in\mathbf{2}$ により，$P\{X\leqq x\}$ は 0 または 1 に等しい．$P\{X\leqq x\}$ は x とともに増大するから，適当な $a\in\mathbf{R}^1$ があって，$P\{X\leqq a-\varepsilon\}=0$, $P\{X\leqq a+\varepsilon\}=1$ $(\varepsilon>0)$．これから $P\{X=a\}=1$ が得られる．

(iv)　$\{X_n(\omega), n=1,2,\cdots\}$ は独立な実確率変数の列とするとき，つぎの諸事象の確率は 0 または 1 であることを示せ．

"$X_n(\omega)$ が $n\to\infty$ のとき収束する",

"$X_n(\omega)$ が有界である",

"$\sum_n |X_n(\omega)|<\infty$",　"$n^{-1}\sum_{k=1}^n X_k(\omega)\longrightarrow 0\quad(n\to\infty)$",

"$X_n(\omega)>0$ i.o.",　"$X_n(\omega)>n$ f.e.".

［ヒント］　$\mathcal{B}_n=\sigma[X_n]$ とおくと，上の諸事象の外延がすべて $\limsup\mathcal{B}_n$ に属することに注意して，Kolmogorov の 0-1 法則を用いよ．

(v)　$X_1(\omega),X_2(\omega),\cdots$ が独立な実確率変数の列であるとき，$Y(\omega)=\limsup_{n\to\infty}X_n(\omega)$ はほとんど確実にある定数（$\pm\infty$ を許す）に等しいことを証明せよ．

［ヒント］　$\mathcal{B}_n=\sigma[X_n]$ としたとき，

$$\{Y(\omega)\leqq y\}\in\limsup_{n\to\infty}\mathcal{B}_n\subset\mathbf{2}.$$

(vi)　$\{\Delta_n\}$ が可分分割の減少列でも $\Delta=\bigwedge_n\Delta_n$ は必ずしも可分でないことを示す反例をあげよ．

［ヒント］　$\Omega=\mathbf{R}^\infty$, $P=\prod\mu_n$（μ_n は連続な 1 次元分布）とし，ω の第 n 座標を $X_n(\omega)$ とすると，$X_1(\omega),X_2(\omega),\cdots$ は独立である．結合変数 $(X_n,X_{n+1},\cdots)$ に対応する分割を Δ_n とすれば，$\{\Delta_n\}$ は可分分割の減少列である．しかし $\Delta=\bigwedge_n\Delta_n$ は可分ではない．もし可分であれば，定理 3.6(ii) により，Δ はある実確率変数 X に対応する．$\Delta\prec\Delta_n$ により，X は $\bigvee_{k\geqq n}\sigma[X_k]$ について可測．したがって定理 3.13 と問(iii)により適当な x に対し $P\{X=x\}=1$. $\{X=x\}\in\Delta=\bigwedge_n\Delta_n$ であるから，$x_1,x_2,\cdots$ があって

$$\{X=x\}=\bigcup_k\{X_k=x_k,\ k\geqq n\}\qquad(\text{定理 } 3.7\,\text{(iii)}),$$

ゆえに

$$P\{X=x\}\leqq\sum_k P\{X_k=x_k,\ k\geqq n\}=0\qquad(\text{矛盾}).$$

§3.5　条件付確率測度

(Ω,P) を基礎の確率空間とし，

$$\Delta=\{\xi_1,\xi_2,\cdots,\xi_n\}$$

を Ω の有限分割とする．要点をわかりやすくするために，

$$P(\xi_i)>0,\qquad i=1,2,\cdots,n$$

と仮定する．おのおのの ξ_i に対して ξ_i の上の確率測度 P_{ξ_i} を

$$(\mathrm{C}_1)\qquad P_{\xi_i}(A)=\frac{P(A)}{P(\xi_i)}\qquad(A\subset\xi_i,\ A\in\mathcal{D}(P))$$

で定義し，これを"ξ_i の中の見本点が出現したことを知った場合における A の**条件付確率**"，または簡単に"ξ_i のもとにおける A の**（条件付）確率**"という．また"確率測度 P_{ξ_i} を ξ_i のもとにおける**（条件付）確率測度**"という．

P^Δ を P から Δ の上に自然に導かれた確率測度とする．すなわち

$$P^\Delta(E)=P\Big(\sum_{\xi\in E}\xi\Big)=\sum_{\xi\in E}P(\xi),\qquad E\subset\Delta.$$

§3.3 で導入した Δ に対応する確率変数

$$X_\Delta(\omega)=\omega \text{ を含む } \Delta \text{ の元}$$

を用いて P^Δ をあらわすと

$$P^{\Delta} = PX_{\Delta}^{-1} = P^{X_{\Delta}}$$

となる．

P^{Δ} も P_{ξ} $(\xi \in \Delta)$ も P から導かれたが，逆に P は P^{Δ} と P_{ξ} によってつぎの式で表わされる．

$$P(A) = \sum_{i=1}^{n} P_{\xi_i}(A \cap \xi_i) P^{\Delta}(\xi_i), \qquad A \in \mathcal{D}(P),$$

積分形でかけば

$$(\mathrm{C}_2) \qquad P(A) = \int_{\Delta} P_{\xi}(A \cap \xi) P^{\Delta}(d\xi), \qquad A \in \mathcal{D}(P).$$

またこの式をみたす ξ の上の確率測度 P_{ξ} があるとせよ．任意の ξ_i を固定し，(C_1) において $A \subset \xi_i$ とすれば，$A \cap \xi_i = A$, $A \cap \xi_j = \phi$ $(j \neq i)$ に注意して，(C_1) を得る．したがって (C_1) と (C_2) とは同等である．これは条件付確率測度 P_{ξ} を (C_2) によって定義してもよいことを示している．しかも (C_2) の式は Δ が一般の分割に対しても意味をもつから，(C_2) によって一般の分割に対応する条件付確率法則を定義できることが予想されるであろう．

(C_2) における P_{ξ} は ξ の上の確率測度であるといったが，もっと精密にいうと ξ の上の完全確率測度である．また (C_2) の A は任意の P 可測集合である．このとき $A \cap \xi$ は一般には P_{ξ} 可測でないから，$P_{\xi}(A \cap \xi)$ が確定しないが，このような例外の ξ 集合は P^{Δ} 零集合であることが証明されるので，(C_2) の右辺の積分は確定した意味をもつのである．

定理 3.18 Δ を可分分割とするとき，すべての $\xi \in \Delta$ に対し，ξ の上の可分完全な確率測度 P_{ξ} を定めて，任意の $A \in \mathcal{D}(P)$ に対し，ほとんどすべて (P^{Δ}) の ξ に対し(例外集合は A に関係する)

$$A \cap \xi \in \mathcal{D}(P_{\xi})$$

で，しかも (C_2) が成立するようにできる．しかもこのような P_{ξ} はほとんど一意である，すなわち上の条件をみたす P_{ξ} が二通りあるとしても，ほとんどすべて (P^{Δ}) の ξ に対し，両者は一致する．

定義 3.1 上の P_{ξ} $(\xi \in \Delta)$ を分割 Δ に対する**条件付確率測度**といい，$P_{\Delta,\xi}$ $(\xi \in \Delta)$ であらわす．

定理 3.18 の証明 P は可分完全であるから，定理 3.1 により $\boldsymbol{R}^1$ のある部分集合の上の K 正則確率測度と強同型である．したがって

$$\Omega \subset \boldsymbol{R}^1, \qquad P \text{ は } \Omega \text{ の上の } K \text{ 正則確率測度}$$

と仮定して一般性を失わない．Δ は可分分割であるから，前述の

$$X_{\Delta}(\omega) = \omega \text{ を含む } \Delta \text{ の元}$$

は Δ 値確率変数である．以後 X_{Δ} を簡単に X とかくことにする．すでにのべたように

$$P^{\Delta} = PX^{-1} = P^{X}$$

で，これは $X(\omega)$ の確率法則であり，したがって Δ の上の可分完全確率測度である．

まず(ほとんど)一意性から証明する．P_{ξ}, P_{ξ}' が (C_2) をみたすとする．A のかわりに $A \cap X^{-1}(E)$ $(E \in \mathcal{D}(P^{\Delta}))$ とおくと，

$$\begin{aligned} P(A \cap X^{-1}(E)) &= \int_{\Delta} P_{\xi}(A \cap X^{-1}(E) \cap \xi) P^{\Delta}(d\xi) \\ &= \int_{\Delta} P_{\xi}'(A \cap X^{-1}(E) \cap \xi) P^{\Delta}(d\xi) \end{aligned}$$

が得られる．$\xi \in E$ か否かに応じて，

$$\xi \subset X^{-1}(E) \quad \text{か} \quad \xi \cap X^{-1}(E) = \phi$$

であるから，上の2積分の積分範囲を E としてよいし，E の上では被積分関数の中の $X^{-1}(E) \cap \xi$ は ξ でおきかえてよい．したがって

$$\int_{E} P_{\xi}(A \cap \xi) P^{\Delta}(d\xi) = \int_{E} P_{\xi}'(A \cap \xi) P^{\Delta}(d\xi)$$

がすべての $E \in \mathcal{D}(P^{\Delta})$ に対してなりたつ．したがって

$$P_{\xi}(A \cap \xi) = P_{\xi}'(A \cap \xi) \ \text{a.e.}(P^{\Delta}) \qquad (\text{a.e.}=\text{almost everywhere}).$$

(この等式は $A \cap \xi$ が $\mathcal{D}(P_{\xi}) \cap \mathcal{D}(P_{\xi}')$ に含まれていることを暗黙のうちに含んでいる.) この等式のなりたたない ξ の全体は A に関係するから，これを $N(A)$ とかくと，$N(A)$ は P^{Δ} 零集合である．上に注意したように $\Omega \subset \boldsymbol{R}^1$ と仮定しているから，Ω の Borel 集合族 $\mathcal{B}(\Omega)$ は Ω 上の乗法族

$$\Omega_r = \Omega \cap (-\infty, r], \qquad r \in \boldsymbol{Q} \quad (\boldsymbol{Q} \text{ は有理数の集合})$$

で生成される．N を $N(\Omega_r)$ $(r \in \boldsymbol{Q})$ の和集合とすると，$\boldsymbol{Q}$ は可算集合であるから，N も P^{Δ} 零集合である．しかも $\xi \in \Delta - N$ に対しては

$$P_\xi(\Omega_r\cap\xi)=P_\xi'(\Omega_r\cap\xi),\qquad r\in\boldsymbol{Q}.$$

$\xi\subset\Omega\subset\boldsymbol{R}^1$ であるから，ξ の上の Borel 集合族 $\mathcal{B}(\xi)$ は $\mathcal{B}(\Omega)\cap\xi$ であり，$\Omega_r\cap\xi$ $(r\in\boldsymbol{Q})$ は $\mathcal{B}(\xi)$ を生成する乗法族であるから，Dynkin 族定理により $\xi\in\Delta-N$ に対し，

$$P_\xi(A)=P_\xi'(A),\qquad A\in\mathcal{B}(\xi).$$

P_ξ, P_ξ' はともに ξ の上の可分完全確率測度であるから，定理 3.3 により，$\xi\in\Delta-N$ に対し P_ξ, P_ξ' はともに ξ の上の正則確率測度である．したがって上の等式から，

$$P_\xi=P_\xi',\qquad \xi\in\Delta-N.$$

これで定理の一意性の部分は証明された．

定理の条件をみたす P_ξ $(\xi\in\Delta)$ の存在を証明しよう．もしこのような P_ξ があるとすれば，上にのべたように

$$P(A\cap X^{-1}(E))=\int_E P_\xi(A\cap\xi)P^\Delta(d\xi),\qquad A\in\mathcal{D}(P),\ E\in\mathcal{D}(P^X).$$

まずこの等式を目標として P_ξ を定めよう．A を固定し

$$Q_A(E)=P(A\cap X^{-1}(E)),\qquad E\in\mathcal{D}(P^\Delta)$$

とおくと，Q_A は Δ の上の有限測度で，しかも

$$\mathcal{D}(Q_A)=\mathcal{D}(P^\Delta),$$
$$P^\Delta(E)=0\Longrightarrow Q_A(E)\leqq P(X^{-1}(E))=P^\Delta(E)=0$$

であるから，測度論で有名な Radon-Nikodym の定理により，P^Δ 可測関数 $f_A(\xi)$ があって

$$(\mathrm{D}_1)\qquad Q_A(E)\ (=P(A\cap X^{-1}(E)))=\int_E f_A(\xi)P^\Delta(d\xi),\qquad E\in\mathcal{D}(P^\Delta)$$

となる．このような $f_A(\xi)$ は上の式から $P_\xi(A\cap\xi)$ の候補と考えられる．

まず個々の A に対し上の等式 (D_1) をみたす $f_A(\xi)$ を任意にとり，$f_A(\xi)$ を P^Δ 測度 0 の ξ 集合の上で修正して $P_\xi(A\cap\xi)$ に到達しようというのが，大体の方針である．この修正が (D_1) をそこなわないことは明らかである．P_ξ は ξ の上の確率測度であるはずであるから，$P_\xi(A\cap\xi)$ は $A\subset\Omega$ の関数として Ω の上の確率測度である．したがって $f_A(\xi)$ も A の関数として Ω の上の確率測度となるようにしなければならない．

まず $0\leqq Q_A(E)\leqq P^\Delta(E)$ に注意して，上の $f_A(\xi)$ は

$$0\leqq f_A(\xi)\leqq 1$$

となるようにとれる．しかも (D_1) において $A=\phi$ または Ω，$E=\Delta$ とおいて

$$\int_\Delta f_\phi(\xi)P^\Delta(d\xi)=P(\phi\cap X^{-1}(\Delta))=0,$$

$$\int_\Delta f_\Omega(\xi)P^\Delta(d\xi)=P(\Omega\cap X^{-1}(\Delta))=P(\Omega)=1,$$

したがって

$$f_\phi(\xi)=0,\qquad f_\Omega(\xi)=1\ \text{a.e.}(P^\Delta)$$

を得るから，これが例外なしになりたつように $f_\phi(\xi), f_\Omega(\xi)$ を修正しておいてよい．

さて $\Omega\subset\boldsymbol{R}^1$ であるから，(D_1) において，$A=\Omega_r\equiv\Omega\cap(-\infty,r]$ とおいて

$$P(\Omega_r\cap X^{-1}(E))=\int_E f_{\Omega_r}(\xi)P^\Delta(d\xi),$$

$f_{\Omega_r}(\xi)$ を $F_\xi(r)$ とかいて

$$(\mathrm{D}_2)\qquad P(\Omega_r\cap X^{-1}(E))=\int_E F_\xi(r)P^\Delta(d\xi),\qquad E\in\mathcal{D}(P^\Delta)$$

が得られる．左辺は r と共に増加するから，$r<r'$ に対し，

$$F_\xi(r)\leqq F_\xi(r')\ \text{a.e.}(P^\Delta).$$

有理数の対 $r<r'$ は可算であるから，“上の不等式がすべての有理数対 $r<r'$ に対してなりたつ”ような ξ の集合の P^Δ 測度は 1 である．したがって例外の ξ 集合の上で $F_\xi(r)$ を修正して，

$$r<r'\ \ (r,r'\in\boldsymbol{Q})\Longrightarrow F_\xi(r)\leqq F_\xi(r')$$

が例外なしになりたつ．修正後も (D_2) はなりたつことは明らかである．以後 r，r' は有理数をあらわすとしておく．$F_\xi(r)$ の r に関する増大性により，

$$F_\xi(r+)=\lim_{n\to\infty}F_\xi(r+n^{-1}),\qquad F_\xi(\pm\infty)=\lim_{n\to\infty}F_\xi(\pm n)$$

は存在し，

$$F_\xi(r+)\geqq F_\xi(r),\qquad 0\leqq F_\xi(-\infty)\leqq F_\xi(\infty)\leqq 1$$

がなりたつ．しかも，$\Omega_{r+n^{-1}}\downarrow\Omega_r$ であるから，

$$\int_\Delta F_\xi(r+)P^\Delta(d\xi)=\lim_{n\to\infty}\int_\Delta F_\xi(r+n^{-1})P^\Delta(d\xi)$$

$$= \lim_{n\to\infty} P(\Omega_{r+n^{-1}}) = P(\Omega_r)$$

$$= \int_\Delta F_\xi(r) P^\Delta(d\xi).$$

これからほとんどすべての ξ に対して

$$F_\xi(r+) = F_\xi(r)$$

が得られる．有理数の集合 $\boldsymbol{Q}$ が可算であるから，上の等式がすべての $r\in\boldsymbol{Q}$ に対してなりたつように $F_\xi(r)$ を修正することができる．同様に $\Omega_{-n}\downarrow\phi$, $\Omega_n\uparrow\Omega$ $(n\to\infty)$ に注意して

$$F_\xi(-\infty) = 0, \quad F_\xi(\infty) = 1$$

となるように $F_\xi(r)$ が修正される．以上を綜合すると，$F_\xi(r)$ を P^Δ 零集合の上で修正して，すべての $\xi\in\Delta$ に対し，$F_\xi(r)$ が $r\in\boldsymbol{Q}$ の増加，右連続な関数で，$-\infty, \infty$ でそれぞれ極限値 $0, 1$ をとるようにし，しかも (D_2) が修正後にもなりたつようにできる．このような $F_\xi(r)$ に対しては1次元分布 $(=\boldsymbol{R}^1$ の上の正則確率測度) μ_ξ が存在して

$$F_\xi(r) = \mu_\xi(-\infty, r], \quad r\in\boldsymbol{Q}$$

となる．したがって

$$P(\Omega_r\cap X^{-1}(E)) = \int_E \mu_\xi(-\infty, r] P^\Delta(d\xi),$$

すなわち

$$P(\Omega\cap(-\infty, r]\cap X^{-1}(E)) = \int_E \mu_\xi(-\infty, r] P^\Delta(d\xi), \quad r\in\boldsymbol{Q}.$$

これからすべての $A\in\mathcal{B}^1$ に対し

$$(\mathrm{D}_3) \qquad P(\Omega\cap A\cap X^{-1}(E)) = \int_E \mu_\xi(A) P^\Delta(d\xi)$$

(この等式は $\mu_\xi(A)$ が ξ に関して P^Δ 可測なことを暗黙のうちに含む) が得られることを示すには，上の両辺が A に関して有限確率測度であり，$A=\boldsymbol{R}^1$, $(-\infty, r]$ $(r\in\boldsymbol{Q})$ に対しては等式がなりたつことに注意して，Dynkin 族定理を適用すればよい．

μ_ξ は $\boldsymbol{R}^1$ の上の正則確率測度であるが，(必要ならば $\mu_\xi(A)$ を修正して) Ω $(\subset \boldsymbol{R}^1)$ の上に集中していることを示そう．P は Ω の上の K 正則確率測度であるから，

コンパクト集合の増大列 $K_n\subset\Omega$ $(n=1, 2, \cdots)$ をとって

$$\lim_{n\to\infty} P(K_n) = P(\Omega) = 1$$

とすることができる．このとき

$$\int_\Delta \lim_{n\to\infty} \mu_\xi(K_n) P^\Delta(d\xi) = \lim_{n\to\infty} \int_\Delta \mu_\xi(K_n) P^\Delta(d\xi) = \lim_{n\to\infty} P(K_n) = 1.$$

左辺の被積分関数は1以下であるから，

$$\lim_{n\to\infty} \mu_\xi(K_n) = 1$$

がほとんどすべての ξ に対してなりたつ．$K_n\subset\Omega$ であるから，μ_ξ の完備性により上の等式は

$$\Omega\in\mathcal{D}(\mu_\xi), \quad \mu_\xi(\Omega) = 1$$

を示している．これの成立しない ξ に対しては μ_ξ を Ω に集中した任意の1次元分布と修正する．この修正後も (D_3) がなりたつ．

μ_ξ は $\boldsymbol{R}^1$ の上の正則確率測度 (したがって K 正則) であるから，$\mu_\xi(\Omega)=1$ により，μ_ξ の Ω の上への制限 ν_ξ は Ω の上の K 正則確率測度である．ν_ξ の定義域 $\mathcal{D}(\nu_\xi)$ は一般に ξ に関係するが，$\mathcal{D}(\mu_\xi)\supset\mathcal{B}^1$, $\mathcal{B}(\Omega)=\mathcal{B}^1\cap\Omega$ により，

$$\mathcal{D}(\nu_\xi) \supset \mathcal{B}(\Omega)$$

となることは確かである．しかも $B\in\mathcal{B}(\Omega)$ は

$$B = B'\cap\Omega, \quad B'\in\mathcal{B}^1$$

とかけるから，

$$\int_E \nu_\xi(B) P^\Delta(d\xi) = \int_E \mu_\xi(B') P^\Delta(d\xi) = P(\Omega\cap B'\cap X^{-1}(E))$$
$$= P(B\cap X^{-1}(E))$$

がすべての $E\in\mathcal{D}(P^\Delta)$ に対してなりたつ．さらに $A\in\mathcal{D}(P)$ に対しても，$\nu_\xi(A)$ がほとんどすべて (P^Δ) の ξ に対して定義され (例外の ξ 集合は A に関係する)，

$$(\mathrm{D}_4) \qquad \int_E \nu_\xi(A) P^\Delta(d\xi) = P(A\cap X^{-1}(E))$$

がなりたつことを証明しよう．P は Ω の上の K 正則確率測度であるから，B_1, $B_2\in\mathcal{B}(\Omega)$ が存在して

$$B_1\subset A\subset B_2, \quad P(B_2-B_1) = 0.$$

$B_1, B_2\in\mathcal{B}(\Omega)\subset\mathcal{D}(\nu_\xi)$ により，$\nu_\xi(B_2-B_1)$ は確定し，

$$\int_{\Delta} \nu_\xi(B_2 - B_1) P^{\Delta}(d\xi) = P(B_2 - B_1) = 0.$$

したがってほとんどすべての ξ に対し

$$\nu_\xi(B_2 - B_1) = 0 \text{ a.e.}(P^{\Delta}),$$

したがって

$$A \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(A) = \nu_\xi(B_1) = \nu_\xi(B_2) \text{ a.e.}(P^{\Delta}).$$

ゆえに

$$\int_E \nu_\xi(A) P^{\Delta}(d\xi) = \int_E \nu_\xi(B_1) P^{\Delta}(d\xi) = P(B_1 \cap X^{-1}(E)) = P(A \cap X^{-1}(E))$$

となり，(D_4) が得られる．

最後にほとんどすべての $\xi \in \Delta$ に対し

$$\xi \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(\xi) = 1$$

であることを証明しよう．もしこれができれば，例外の ξ 集合の上では ν_ξ を上の条件がなりたつように修正（例えば ξ の中の1点 ω_0 をとり $\nu_\xi = \delta_{\omega_0}$ とおく）して，上の条件が例外なしになりたつことになり，しかも修正後も (D_4) はそこなわれない．したがって，ν_ξ の ξ の上への制限を P_ξ とすれば，$\nu_\xi(A) = P_\xi(A \cap \xi)$ となり，(D_4) で $E = \Delta$ とおいて P_ξ が (C_2) をみたすことがわかる．（ν_ξ の K 正則性により，P_ξ も K 正則，したがって可分完全であることに注意せよ．）

残る所は $\xi \in \mathcal{D}(\nu_\xi)$，$\nu_\xi(\xi) = 1$ の証明である．Δ は可分分割であるから，可算族 $\{A_n\} \subset \mathcal{D}(P)$ で生成される．明らかに $A_n = X^{-1}(E_n)$，$E_n \in \mathcal{D}(P^{\Delta})$ とかける．(D_4) で $A = A_n$，$E = E_n$ とおくと，

$$\int_{E_n} \nu_\xi(A_n) P^{\Delta}(d\xi) = P(A_n \cap X^{-1}(E_n)) = P(X^{-1}(E_n)) = P^{\Delta}(E_n).$$

$\nu_\xi(A) \leqq 1$ であるから，E_n の上のほとんど到るところ，

$$A_n \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(A_n) = 1.$$

同様に (D_4) で $A = A_n^c$，$E = E_n^c$ とおいて，E_n^c の上でほとんど到るところ，

$$A_n^c \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(A_n^c) = 1.$$

上にあらわれる例外の ξ 集合を N_n, N_n' とし，

$$N = \bigcup_n (N_n \cup N_n')$$

とすれば，$P^{\Delta}(N) = 0$ で，しかも $\xi \in \Delta - N$ では，すべての n に対し

$$\xi \in E_n \Longrightarrow A_n \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(A_n) = 1,$$
$$\xi \in E_n^c \Longrightarrow A_n^c \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(A_n^c) = 1$$

がなりたつ．任意の A_n に対し $A_n^1 = A_n$，$A_n^0 = A_n^c$ とし，E_n に対しても同様の記号を用いると，上の事実は，$\xi \in \Delta - N$ に対して

$$\xi \in E_n^{\varepsilon} \Longrightarrow A_n^{\varepsilon} \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(A_n^{\varepsilon}) = 1, \qquad n = 1, 2, \cdots$$

を意味する．したがって

$$\xi \in \bigcap_n E_n^{\varepsilon_n} \Longrightarrow \bigcap_n A_n^{\varepsilon_n} \in \mathcal{D}(\nu_\xi), \quad \nu_\xi\Big(\bigcap_n A_n^{\varepsilon_n}\Big) = 1.$$

$\{A_n\} = \{X^{-1}(E_n)\}$ が Δ を生成するから，上の事実により，$\xi \in \Delta - N$ に対し，

$$\xi \in \bigcap_n E_n^{\varepsilon_n} \Longrightarrow \xi = \bigcap_n A_n^{\varepsilon_n} \Longrightarrow \xi \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(\xi) = 1.$$

$\xi \in \Delta - N$ に対し，

$$\xi \in E_n \text{ か否かに応じて} \quad \varepsilon_n = 1, 0$$

とすれば，当然 $\xi \in \bigcap_n E_n^{\varepsilon_n}$ となるから，結局すべての $\xi \in \Delta - N$ に対し

$$\xi \in \mathcal{D}(\nu_\xi), \quad \nu_\xi(\xi) = 1$$

が証明されたことになる．$P(N) = 0$ であるから，定理は完全に証明された．∎

註1　定理3.1より，P が $\boldsymbol{R}^1$ のある部分集合の上の K 正則確率測度に強同型であるから，$\Omega \subset \boldsymbol{R}^1$ として証明をすすめたが，この強同型を示す全単射の逆写像でもとにもどると，それに応じて ξ の上の可分完全確率測度も同様な確率測度に移ることに注意すべきである．

註2　$P_{\Delta,\xi}$ は ξ の上の可分完全確率測度であるが，これを Ω の上へ拡張したものを再び $P_{\Delta,\xi}$ とかくと，これは Ω の上の可分完全確率測度で，しかも $P_{\Delta,\xi}(\xi) = 1$ となっている．もとの $P_{\Delta,\xi}$ はこの新しい $P_{\Delta,\xi}$ の ξ への制限である．この新しい $P_{\Delta,\xi}$ を特長づける条件は

$$(\mathrm{C}_3) \qquad P(A) = \int_{\Delta} P_{\Delta,\xi}(A) P^{\Delta}(d\xi), \quad P_{\Delta,\xi}(\xi) = 1$$

となる．すべての $A \in \mathcal{D}(P)$ に対し '$A \in \mathcal{D}(P_{\Delta,\xi})$ a.e.(P^{Δ})' となることは前と同様である．

註3　(C_3) のかわりに

$$(\mathrm{C}_3') \qquad P(A \cap X_{\Delta}^{-1}(E)) = \int_E P_{\Delta,\xi}(A) P^{\Delta}(d\xi), \quad E \in \mathcal{D}(P^{\Delta})$$

といってもほとんど同等である．実際，(C_3) から (C_3') がでることは

$$\begin{aligned}\xi\in E &\Longrightarrow \xi\subset X_\Delta^{-1}(E)\\ &\Longrightarrow P_{\Delta,\xi}(B\cap X_\Delta^{-1}(E)) = P_{\Delta,\xi}(B\cap X_\Delta^{-1}(E)\cap\xi) = P_{\Delta,\xi}(B\cap\xi)\\ &\qquad\qquad = P_{\Delta,\xi}(B),\\ \xi\in E^c &\Longrightarrow \xi\cap X_\Delta^{-1}(E)=\phi\\ &\Longrightarrow P_{\Delta,\xi}(B\cap X_\Delta^{-1}(E)) = P_{\Delta,\xi}(B\cap X_\Delta^{-1}(E)\cap\xi) = P_{\Delta,\xi}(\phi)=0\end{aligned}$$

を用いて容易に示される．また (C_3') から (C_3) の第1式をだすには，(C_3') で $E=\Delta$ とおけばよい．(C_3') の E のかわりに Δ の2点を分離する A_n $(n=1,2,\cdots)$ をいれて，定理3.18の証明で $\nu_\xi(\xi)=1$ a.e.(P^Δ) を導いたのと同様に‘$P_{\Delta,\xi}(\xi)=1$ a.e.(P^Δ)’が導ける．したがって (C_3) の方がより精密ないい方といえるが，その差は P^Δ 零集合の上だけのもので，本質的な差はない．

註4　$X(\omega)$ が S 値確率変数としたとき，

$$\Delta_X = \{X^{-1}(s)\,|\,s\in X(\Omega)\}$$

は Ω の可分分割であるから，これに対して $P_{\Delta_X,X^{-1}(s)}$ が定義される．$X^{-1}(s)=\{\omega\,|\,X(\omega)=s\}$ であることを念頭において，$P_{\Delta_X,X^{-1}(s)}$ を $P_{X=s}$ とかく．$P_{X=s}$ を‘$X=s$ のもとにおける**条件付確率法則**’という．とくに $P\{X(\omega)=s\}>0$ のときには

$$P_{X=s}(A) = \frac{P(\{X=s\}\cap A)}{P\{X=s\}},\qquad A\in\mathcal{D}(P)$$

となることは容易にわかる．この際 (C_3') に相当する式は

$$(C_3'')\qquad P(A\cap X^{-1}(E)) = \int_E P_{X=s}(A)P^X(ds)$$

となる．

註5　$P^\Delta=PX_\Delta^{-1}$ を用い (C_3') の右辺を積分変換公式によりかきかえて

$$\begin{aligned}P(A\cap X_\Delta^{-1}(E)) &= \int_{X_\Delta^{-1}(E)} P_{\Delta,X_\Delta(\omega)}(A)P(d\omega)\\ &= E(P_{\Delta,X_\Delta(\omega)}(A),X_\Delta^{-1}(E)),\qquad E\in\mathcal{D}(P^{X_\Delta})\end{aligned}$$

とし，さらに Δ で定まる σ 加法族 $\mathcal{B}_\Delta$ の定義

$$\mathcal{B}_\Delta = X_\Delta^{-1}(\mathcal{D}(P^{X_\Delta})) = \bar{\sigma}[X_\Delta]$$

を思いおこすと，上の式は

$$P(A\cap B) = E(P_{\Delta,X_\Delta(\omega)}(A),B),\qquad B\in\mathcal{B}_\Delta$$

となる．$A\in\mathcal{D}(P)$ を固定したとき $P_{\Delta,\xi}(A)$ はほとんどすべての ξ に対して定義されているにすぎないが，P^Δ $(=P^{X_\Delta})$ 可測関数とほとんど (P^Δ) 一致している．したがって $P_{\Delta,X_\Delta(\omega)}(A)$ はほとんどすべて (P) の ω に対して定義され，$\mathcal{B}_\Delta$ $(=\bar{\sigma}[X_\Delta])$ に関して可測な関数とほとんど一致している．この事実を

$$P_{\Delta,X_\Delta(\omega)}(A)\in\mathcal{B}_\Delta \text{ a.s.}$$

とかくことにしよう．かくして $\varphi_\omega(A)=P_{\Delta,X_\Delta(\omega)}(A)$ を定める条件を $\mathcal{B}_\Delta$ だけの言葉で

$$P(A\cap B) = E(\varphi_\omega(A),B),\qquad B\in\mathcal{B}_\Delta,\ \varphi_\omega(A)\in\mathcal{B}_\Delta \text{ a.s.}$$

とかいてもよいことがわかるであろう．この意味で

$$\varphi_\omega = P_{\Delta,X_\Delta(\omega)}$$

を $\mathcal{B}_\Delta$ **のもとにおける条件付確率測度**といい，$P_{\mathcal{B}_\Delta,\omega}$ とかく．$P_{\mathcal{B}_\Delta,\omega}$ の ω を省略して $P_{\mathcal{B}_\Delta}$ とかくことが多いが，これが ω に関係していることを忘れてはならない．かくしてすべての $A\in\mathcal{D}(P)$ に対し，

$$P(A\cap B) = E(P_{\mathcal{B}_\Delta}(A),B)\quad (B\in\mathcal{B}_\Delta),\qquad P_{\mathcal{B}_\Delta}(A)\in\mathcal{B}_\Delta \text{ a.s.}$$

を得る．この事実を背景として，一般の σ 加法族 $\mathcal{B}\subset\mathcal{D}(P)$ に対して $\mathcal{B}$ のもとにおける条件付確率測度 $P_{\mathcal{B}}=P_{\mathcal{B},\omega}$ をつぎの3条件で定義しよう．

(C.1)　個々の $\omega\in\Omega$ に対し，$P_{\mathcal{B},\omega}$ は Ω の上の可分完全確率測度である，

(C.2)　すべての $A\in\mathcal{D}(P)$ に対し，$P_{\mathcal{B},\omega}(A)$ はほとんど確実に確定し，
$P_{\mathcal{B},\omega}(A)\in\mathcal{B}$ a.s.，

(C.3)　$P(A\cap B) = E(P_{\mathcal{B}}(A),B)$　$(B\in\mathcal{B})$．

$P_{\mathcal{B},\omega}$ の存在をいうには，

$$P(A\cap C)\leqq P(C),\qquad C\in\mathcal{B}$$

に注意して Radon-Nikodym の定理により

$$P(A\cap C) = \int_C f_A(\omega)P(d\omega),\qquad f_A\in\mathcal{B}$$

とかき，定理3.18の証明にならって $f_A(\omega)$ を変形して $P_{\mathcal{B}}=P_{\mathcal{B},\omega}$ を定めたらよい．$P_{\mathcal{B}}$ がほとんど一意であることの証明も定理3.18の証明と同様である．$P_{\Delta,\xi}$ のときには

$$P_{\Delta,\xi}(\xi)=1$$

がいえたが，$P_{\mathcal{B}}$ の場合にはこれに対応することは一般にはいえない．これにほぼ近いこととしては

$$B \in \mathcal{B} \Longrightarrow P_{\mathcal{B},\omega}(B) = 1_B(\omega) \text{ a.s.}$$

である．これは

$$P(B) = P(B \cap B) = E(P_{\mathcal{B}}(B), B),$$
$$0 = P(B \cap B^c) = E(P_{\mathcal{B}}(B), B^c)$$

から直ちに得られる．

とくに $\mathcal{B}=\mathcal{B}_\Delta$ (Δ は可分分割) のときには，上の定義のしかたをよく見ると，$P_{\mathcal{B}_\Delta,\omega}$ は ω が Δ のおのおのの元の上では一定となるようにできる．したがってこれを $P_{\Delta,\xi}$ とかいてよい．すなわち

$$P_{\Delta,\xi} = P_{\mathcal{B}_\Delta,\omega}, \qquad \omega \in \xi.$$

$B \in \mathcal{B}_\Delta$ は $X_\Delta^{-1}(E)$ ($E \in \mathcal{D}(P^\Delta)$) の形にかけるから，上に示した式により，

$$E \in \mathcal{D}(P^\Delta) \Longrightarrow P_{\Delta,\xi}(X_\Delta^{-1}(E)) = 1_E(\xi) \text{ a.e.}(P^\Delta).$$

Δ の2点を分離する可算族 $\{E_n\} \subset \mathcal{D}(P^\Delta)$ をとると，

$$P_{\Delta,\xi}(X_\Delta^{-1}(E_n)) = 1_{E_n}(\xi), \qquad P_{\Delta,\xi}(X_\Delta^{-1}(E_n{}^c)) = 1_{E_n{}^c}(\xi),$$
$$n = 1, 2, \cdots, \text{ a.e.}(P^\Delta).$$

ここで例外の ξ 集合 N は n に無関係にとれる．例外の $\xi \in N$ に対しては $P_{\Delta,\xi}$ を ξ の上に集中した，Ω の上の任意の可分完全確率測度 (たとえば ξ の中の1点 ω_0 における δ 測度 δ_{ω_0}) とおけば，上の式は例外なしになりたつ．Δ の任意の点 ξ をとり，$\xi \in E_n$，$\xi \notin E_n$ に応じて $E_n'=E_n$，$E_n{}^c$ とおくと，

$$P_{\Delta,\xi}(X_\Delta^{-1}(E_n')) = 1, \qquad n = 1, 2, \cdots,$$

$$P_{\Delta,\xi}(\xi) = P_{\Delta,\xi}\left(X_\Delta^{-1}\left(\bigcap_n E_n'\right)\right) = P_{\Delta,\xi}\left(\bigcap_n X_\Delta^{-1}(E_n')\right) = 1.$$

このようにして $P_{\mathcal{B},\omega}$ から $P_{\Delta,\xi}$ が得られ，さらにこれから $P_{X=x}$ が得られることは，前述の通りである．

例題 3.5 (i) $\Omega=[0,1]^2$，P を Ω の上の Lebesgue 測度とする．Δ を '$x+y=\xi$ ($0 \leqq \xi \leqq 2$) の形の直線と Ω との交わり' の全体としたとき，$P_{\Delta,\xi}$ を定めよ (Δ の点は $\xi \in [0,2]$ で代表される)．

(ii) $\Omega=\boldsymbol{R}^2$, P を密度 $f(x,y)$ の2次元分布とする．$\omega=(x,y)$ に対して $X(\omega)=x$，$Y(\omega)=y$ とする．$X=x$ の条件のもとにおける Y の確率法則 $P_{X=x}{}^Y$ は

$$\frac{f(x,y)}{\displaystyle\int_{R^1} f(x,y)\,dy} \qquad \text{(分母の積分が0でないとする)}$$

を密度とする1次元分布となることを示せ．分母の積分が0となるときには，どう理解したらよいかを考えよ．

(iii) $\mathcal{B}_1, \mathcal{B}_2$ が独立のとき，$B_2 \in \mathcal{B}_2$ に対し

$$P_{\mathcal{B}_1}(B_2) = P(B_2) \text{ a.s.}$$

を証明せよ．

[ヒント] $B_1 \in \mathcal{B}_1$ に対し，$E(P_{\mathcal{B}_1}(B_2), B_1) = P(B_2 \cap B_1) = P(B_2)P(B_1) = E(P(B_2), B_1)$．

(iv) $P_{\mathcal{D}(P)}(B) = 1_B(\omega)$ a.s. を証明せよ．

[ヒント] (iii) と同様．

§3.6 条件付確率測度の性質

前節において3種類の条件付確率測度

$$P_\Delta = P_{\Delta,\xi}, \qquad P_{X=x}, \qquad P_{\mathcal{B}} = P_{\mathcal{B},\omega}$$

についてのべ，それぞれ ξ, x, ω に依存する Ω の上の可分完全確率測度であることを示した．Δ に対応する確率変数 X_Δ，X による分割 Δ_X，すなわち

$$X_\Delta(\omega) = \omega \text{ を含む } \Delta \text{ の元}, \qquad \Delta_X = \{X^{-1}(s) \mid s \in X(\Omega)\}$$

を考えると，

$$P_{\Delta,\xi} = P_{X_\Delta=\xi}, \qquad P_{X=s} = P_{\Delta_X, X^{-1}(s)}$$

となるから，$P_{X=s}$ の性質を調べておけば，$P_{\Delta,\xi}$ の性質はそれから容易に導かれる．しかし $P_{\mathcal{B}}=P_{\mathcal{B},\omega}$ はもう少し一般的であって，とくに $\mathcal{B}=\mathcal{B}_\Delta$ または $\mathcal{B}=\bar{\sigma}[X] \equiv X^{-1}(\mathcal{D}(P^X))$ のときだけが，$P_{\Delta,\xi}, P_{X=s}$ に対応するにすぎない．

定理 3.19 $X(\omega), Y(\omega)$ をそれぞれ S, T の中の値をとる確率変数とするとき，

(i) $$P\{X(\omega) \in A, Y(\omega) \in B\} = \int_A P^X(ds) P_{X=s}\{Y \in B\}$$
$$(A \in \mathcal{D}(P^X),\ B \in \mathcal{D}(P^Y)),$$

(ii) $$P\{(X(\omega), Y(\omega)) \in C\} = \int_S P^X(ds) P_{X=s}\{Y \in C(s)\}$$
$$(C \in \mathcal{D}(P^{(X,Y)}),\ C(s) = \{t \in T \mid (s,t) \in C\}).$$

証明　$Y^{-1}(B) \in \mathcal{D}(P)$ であるから，定理 3.18 で $\Delta=\Delta_X$ とおいて，ほとんどすべて (P^X) の s に対し

$$Y^{-1}(B) \cap \{X=s\} \in \mathcal{D}(P_{X=s}),$$

$$P\{X^{-1}(A) \cap Y^{-1}(B)\} = \int_A P^X(ds) P_{X=s}\{Y^{-1}(B)\} \qquad (A \in \mathcal{D}(P^X))$$

が得られる．これは定理の第1式である．

$X(\omega)$, $Y(\omega)$ は確率変数であるから，S, T の上の可算分離族 $\{A_n\} \subset \mathcal{D}(P^X)$, $\{B_n\} \subset \mathcal{D}(P^Y)$ が存在する．C が

$$A_m \times B_n, \qquad m, n = 1, 2, \cdots$$

の形の集合のときには，すでに証明された第1式により，第2式は当然成立する．このような形の集合の全体を $\mathcal{C}$ とすると，Dynkin 族定理を用いて，すべての $C \in \sigma[\mathcal{C}]$ に対して第2式が成立することもわかる．§3.2 の結合変数の定義でのべたように $P^{(X,Y)}$ はその $\sigma[\mathcal{C}]$ への制限 $P^{(X,Y)}|_{\sigma[\mathcal{C}]}$ の Lebesgue 拡大と一致する．したがって $C \in \mathcal{D}(P^{(X,Y)})$ に対し，$C_1, C_2 \in \sigma[\mathcal{C}]$ が存在して

$$C_1 \subset C \subset C_2, \qquad P^{(X,Y)}(C_2 - C_1) = 0.$$

C_1, C_2 に対しては第2式が成立するから，

$$C_1(s) \subset C(s) \subset C_2(s)$$

に注意して，第2式が C に対しても成立することがいえる．▌

註1　上の定理の第2式は $P_{X=s}\{Y \in C(s)\}$ がほとんどすべての (P^X) に対して定義され，しかも

$$P_{X=s}(Y \in C(s))\ (s\ \text{の関数として}) \in \mathcal{D}(P^X)\ \text{a.e.}(P^X).$$

すなわち，この s の関数がある P^X 可測関数とほとんど (P^X) 一致するということも含まれていることに注意すべきである．

註2　$P_{X=s}$ は Ω の上の可分完全確率測度であるから，$(\Omega, P_{X=s})$ は可分完全確率空間である．上に注意したように，ほとんどすべての (P^X) の s に対し

$$Y^{-1}(B_n) \in \mathcal{D}(P_{X=s}), \qquad n = 1, 2, \cdots$$

がなりたつ．ここで例外の s 集合 $N(P^X(N)=0)$ は n に無関係にとれる．したがって $s \in S-N$ に対し $Y(\omega)$ は $(\Omega, P_{X=s})$ の上の T 値確率変数となるから，その確率法則が考えられる．これを $X=s$ のもとにおける Y の**条件付確率法則**といい，$P_{X=s}{}^Y$ であらわす．この記号を用いると定理の第2式は

$$P^{(X,Y)}(C) = \int_S P^X(ds) P_{X=s}{}^Y(C(s))$$

とかける．——

$Y(\omega)$ が位相確率変数のときには，それを $(\Omega, P_{X=s})$ の上で考えてもほとんどすべて (P^X) の s に対し位相確率変数となることは容易にわかる．とくに $Y(\omega)$ が (Ω, P) の上で一般化された実確率変数のときには，$(\Omega, P_{X=s})$ の上で考えても同様で，その期待値 $E_{X=s}(Y)$（**条件付期待値**または**条件付平均値**という）を定義することができる．(P^X 測度0の例外があることはいうまでもない.) $E_{X=s}(Y, B)$ についても同様である．

定理 3.19 と積分の定義を用いて，直積測度に関する Fubini の定理に類似の次の定理が得られる．

定理 3.20　$X(\omega)$, $Y(\omega)$ をそれぞれ S, T の中の値をとる確率変数とするとき，

$$E\{f(X(\omega), Y(\omega))\} = \int_S P^X(ds) E_{X=s}(f(s, Y))$$

がなりたつ．正確にいうと，左辺が確定すれば右辺が確定し，上の等式がなりたつ．——

一般の σ 加法族 $\mathcal{B} \subset \mathcal{D}(P)$ に対しても $P_{\mathcal{B},\omega}$ から $P_{\mathcal{B},\omega}{}^Y$, $E_{\mathcal{B},\omega}(Y)$ が定義できる．

例題 3.6　(i)　$E(f(X, Y))$ が確定するとき

$$E_{X=s}(f(X, Y)) = E_{X=s}(f(s, Y))\ \text{a.e.}(P^X)$$

を示せ．

［ヒント］　$P_{X=s}$ の定義に注意せよ．

(ii)　$X(\omega)$, $Y(\omega)$ が独立のとき，

$$P_{X=s}{}^Y = P^Y\ \text{a.e.}(P^X),$$

$$E_{X=s}(f(X, Y)) = E(f(s, Y))\ \text{a.e.}(P^X)$$

$$(E(f(X, Y))\ \text{が確定すると仮定})$$

を示せ．

(iii)　$X(\omega)$, $Y(\omega)$, $Z(\omega)$ がそれぞれ S, T, U の中の値をとる確率変数とする．$E(f(X, Y, Z))$ が確定すれば

$$E(f(X, Y, Z)) = \int_S P^X(ds) \int_T P_{X=s}{}^Y(dt) E_{(X,Y)=(s,t)}(f(s, t, Z))$$

がなりたつことを示せ.

§3.7　実確率変数

基礎の確率空間 (Ω, P) の上のすべての実確率変数の全体を $L^0=L^0(\Omega, P)$ であらわす. L^0 の中の2元 X, Y は, $X(\omega)=Y(\omega)$ a.s. のとき, しばしば同一視される. L^0 の中に距離

$$\rho_0(X, Y) = E(|X-Y|\wedge 1) \qquad (a\wedge b=\min(a, b))$$

を入れる. これが距離の性質

$$\rho_0(X, X) = 0, \quad \rho_0(X, Y) = \rho_0(Y, X),$$
$$\rho_0(X, Y)+\rho_0(Y, Z) \geqq \rho_0(X, Z)$$

をみたすことは明らかであるが, さらにつぎの意味で ρ_0 は, 分離公理をみたしている.

$$\rho_0(X, Y) = 0 \Longrightarrow X(\omega) = Y(\omega) \ \text{a.s.}$$

$X_1, X_2, \cdots, X\in L^0$ とする. 任意の $\varepsilon>0$ に対し

$$P\{|X_n-X|>\varepsilon\} \longrightarrow 0 \qquad (n\to\infty)$$

のとき, X_n は X に**確率収束**するといい,

$$X_n \longrightarrow X \ \text{i.p.} \qquad (\text{i.p.}=\text{in probability})$$

であらわす. これは

$$\rho_0(X_n, X) \longrightarrow 0$$

と同等であることはつぎの不等式から明らかである.

$$\varepsilon P\{|X-Y|>\varepsilon\} \leqq \rho_0(X, Y) \leqq \varepsilon+P\{|X-Y|>\varepsilon\} \qquad (0<\varepsilon<1).$$

実際, $0<\varepsilon<1$ により,

$$\rho_0(X, Y) = E(|X-Y|\wedge 1) \geqq E(\varepsilon, |X-Y|>\varepsilon) \geqq \varepsilon P\{|X-Y|>\varepsilon\},$$
$$\rho_0(X, Y) \leqq E(\varepsilon, |X-Y|\leqq\varepsilon)+E(1, |X-Y|>\varepsilon) \leqq \varepsilon+P\{|X-Y|>\varepsilon\}.$$

$X_n\to X$ a.s. のとき, X_n は X に**概収束**するという. 概収束ならば確率収束であることは明らかであるが, 逆は必ずしも真ではない. 例えば

$$\Omega = [0, 1), \qquad P = \Omega \text{ の上の Lebesgue 測度},$$

$$X_{n,k}(\omega) = \text{区間}\left[\frac{k-1}{n}, \frac{k}{n}\right) \text{の指示関数}$$

としたとき, $X_{1,1}, X_{2,1}, X_{2,2}, X_{3,1}, X_{3,2}, X_{3,3}, \cdots$ を $X_1, X_2, \cdots$ とかくと, X_n は X

$\equiv 0$ に確率収束するが, 概収束はしない.

'$X_n\to X$ i.p.' でも '$X_n\to X$ a.s.' とはいえないが, $\{X_n\}$ のある部分列は X に概収束する. 実際 '$X_n\to X$ i.p.' の仮定から, $\{X_n\}$ の部分列 $\{Y_n\}$ をとって

$$P\{|Y_n-X|>2^{-n}\} < 2^{-n}$$

とすることができる. Borel-Cantelli の補題により, ほとんど確実に, ある番号以後

$$|Y_n-X| \leqq 2^{-n}$$

がなりたつ. これは

$$Y_n \longrightarrow X \ \text{a.s.}$$

を意味する.

定理 3.21　(L^0, ρ_0) は完備可分距離空間である.

証明　$\{X_n\}$ が ρ_0 に関して Cauchy 列とする. すなわち

$$\rho_0(X_m, X_n) \longrightarrow 0 \qquad (m, n\to\infty)$$

とする. $\{X_n\}$ の部分列 $\{Y_n\}$ を適当にとって

$$\rho_0(Y_{n+1}, Y_n) < 2^{-n}, \qquad n = 1, 2, \cdots$$

とすることができる. これから

$$E\left\{\sum_n |Y_{n+1}-Y_n|\wedge 1\right\} = \sum_n \rho(Y_{n+1}, Y_n) < \infty,$$
$$\sum_n (|Y_{n+1}-Y_n|\wedge 1) < \infty \ \text{a.s.}$$

が得られる. したがって

$$|Y_{n+1}-Y_n|\wedge 1 \longrightarrow 0 \ \text{a.s.}$$

これはある番号以後

$$|Y_{n+1}-Y_n|\wedge 1 = |Y_{n+1}-Y_n|$$

となることを意味する. したがって上の不等式から

$$\sum_n |Y_{n+1}-Y_n| < \infty \ \text{a.s.}$$

が得られる. これから Y_n がある $Y\in L^0$ に概収束することがわかる. したがって

$$\rho_0(Y_n, Y) \longrightarrow 0 \qquad (n\to\infty).$$

$\{Y_n\}$ は $\{X_n\}$ の部分列であるから, $\rho_0(X_m, X_n)\to 0$ $(m, n\to\infty)$ により

$$\rho_0(X_m, Y_m) \longrightarrow 0 \qquad (m\to\infty).$$

上の2式から

$$\rho_0(X_n, Y) \longrightarrow 0 \qquad (n\to\infty)$$

がでる．これで (L^0, ρ_0) の完備性が証明された．

可分性は P が $\boldsymbol{R}^1$ のある部分集合 S の上の K 正則確率測度 μ と強同型なことに注意して，$(\Omega, P)=(S,\mu)$ の場合に証明すればよい．このときには

$$\sum_{i=1}^n a_i 1_{[\alpha_i,\beta_i]}, \qquad a_i, \alpha_i, \beta_i \in \boldsymbol{Q}$$

の形の関数の全体は (L^0,ρ_0) の中で稠密であり，しかもこのような関数は可算個しかない．∎

L^0 の部分空間

$$L^p = L^p(\Omega,P) = \{X\in L^0 \mid E(|X|^p)<\infty\} \qquad (1\leqq p<\infty)$$

はノルム

$$\|X\|_p = (E(|X|^p))^{1/p}$$

に関して Banach 空間であることは測度論でよく知られているが，L^0 の可分性と同様に L^p の可分性を証明することができる．

X の確率法則 P^X は1次元分布であるから，定理2.12により $(|M|_p(P^X))^{1/p}$ は p とともに増大する．積分の変換公式により，

$$|M|_p(P^X) = \int_{R^1}|x|^pP^X(dx) = \int_\Omega |X(\omega)|^pP(d\omega) = E(|X|^p)$$

であるから，$\|X\|_p=(E(|X|^p))^{1/p}$ も p と共に増大する．(この事実は定理2.12の証明にならって直接証明することもできる.) したがって，$p<q$ のときには，

$$L^q\subset L^p,$$
$$\|X_n-X\|_q \longrightarrow 0 \Longrightarrow \|X_n-X\|_p\longrightarrow 0$$

がなりたつ．また

$$E(|X|^p) \geqq E(|X|^p, |X|\geqq\varepsilon) \geqq \varepsilon^pP(|X|\geqq\varepsilon)$$

により，

$$\|X_n-X\|_p\longrightarrow 0 \Longrightarrow X_n\longrightarrow X \text{ i.p.}$$

が得られる．$\|X_n-X\|_p\to0$ のとき，X_n は X に **p 次平均収束**するという．p 次平均収束列が概収束部分列をもつことは，p 次平均収束から確率収束がでること

により明らかである．

X_n の確率法則が X の確率法則に(§2.5 b)の意味で)収束するとき，X_n は X に**法則収束**するという．これは $X_1, X_2, \cdots, X$ の定義されている確率空間が違っていても考えられる収束概念である．任意の実確率変数 Y に対して，積分変換公式により

$$E(f(Y)) = \int_{R^1} f(y)P^Y(dy)$$

であるから，定理2.10により，"X_n が X に法則収束する" ためには "コンパクトな台をもつ任意の連続関数 g に対し

$$E(g(X_n)) \longrightarrow E(g(X))$$

となる" ことが必要十分である．

定理 3.22 '$X_n\to X$ i.p.' ならば，X_n は X に法則収束する．

証明 g をコンパクトな台をもつ $\boldsymbol{R}^1$ の上の連続実関数とすると，g は一様連続かつ有界である．$A=\sup|g(x)|$ とする．

$$\begin{aligned} |E(g(X_n))-E(g(X))| &\leqq E(|g(X_n)-g(X)|)\\ &\leqq E(|g(X_n)-g(X)|, |X_n-X|\leqq\varepsilon)+2AP\{|X_n-X|>\varepsilon\}\\ &\leqq \sup\{|g(x)-g(y)|\mid |x-y|<\varepsilon\}+2AP\{|X_n-X|>\varepsilon\}.\end{aligned}$$

'$X_n\to X$ i.p.' により

$$\begin{aligned}\limsup_{n\to\infty}|E(g(X_n))-E(g(X))| &\leqq \sup\{|g(x)-g(y)|\mid|x-y|<\varepsilon\}\\ &\longrightarrow 0 \qquad (\varepsilon\to0).\end{aligned}$$

これから $E(g(X_n))\to E(g(X))$ $(n\to\infty)$ がでるから，X_n は X に法則収束する．∎

以上の諸種の収束概念の関係を整理すると，つぎのようになる．

$$\begin{tikzcd}[row sep=small] \text{概収束} \arrow[dr, Rightarrow] & & \\ & \text{確率収束} \arrow[r, Rightarrow] & \text{法則収束}. \\ \text{平均収束} \arrow[ur, Rightarrow] & & \end{tikzcd}$$

1次平均収束すれば，確率収束するから，適当な部分列が概収束する．$X_n\in L^1$ $(n=1,2,\cdots)$ が X に概収束しても，$X\in L^1$ とは限らないし，まして X_n が X に1次平均収束するとはいえない．では $\{X_n\}$ にいかなる条件をつければ，'$X_n\to X$ a.s.' から '$\|X_n-X\|_1\to0$' がでるであろうか．これに答えるために一様可積分という概念を導入する．確率変数の系

$$\{X_\lambda, \lambda \in \Lambda\} \subset L^1$$

があって

$$\sup_\lambda E(|X_\lambda|, |X_\lambda| > A) \longrightarrow 0 \qquad (A \to \infty)$$

のとき，$\{X_\lambda, \lambda \in \Lambda\}$ は**一様可積分**という．$X_\lambda \in L^1$ であるから，個々の λ に対しては

$$E(|X_\lambda|, |X_\lambda| > A) \longrightarrow 0 \qquad (A \to \infty)$$

となることは明らかであるが，一様可積分性はこの収束が $\lambda \in \Lambda$ について一様なことを意味する．

定理 3.23　$\{X_\lambda, \lambda \in \Lambda\}$ が一様可積分ならば，

(i)　$\|X_\lambda\|_1$ $(\lambda \in \Lambda)$ は有界である，

(ii)　$\sup_\lambda P\{|X_\lambda| > A\} \longrightarrow 0 \quad (A \to \infty)$，

(iii)　$P(B) \longrightarrow 0 \Longrightarrow \sup_\lambda E(|X_\lambda|, B) \longrightarrow 0$.

註　'(ii)＋(iii)' から $\{X_\lambda\}$ の一様可積分がでることは，(iii) において，B のかわりに $B_\lambda = \{|X_\lambda| > A\}$ とおいて見れば，すぐにわかる．したがって '(ii)＋(iii)' は一様可積分のための必要十分条件である．下の証明からわかるように '(i) $\Longrightarrow$ (ii)' であるから '(i)＋(iii)' も同様である．

定理 3.23 の証明　(i) は

$$\begin{aligned}\|X_\lambda\|_1 &\leqq E(|X_\lambda|, |X_\lambda| > A) + E(|X_\lambda|, |X_\lambda| \leqq A)\\ &\leqq E(|X_\lambda|, |X_\lambda| > A) + A.\end{aligned}$$

一様可積分の仮定により，A を λ に無関係に十分大きくとって上の第1項を1以下にできるから，$\|X_\lambda\|_1$ $(\lambda \in \Lambda)$ は有界である．

(ii)　上の (i) と

$$\|X_\lambda\|_1 \geqq E(|X_\lambda|, |X_\lambda| > A) \geqq AP\{|X_\lambda| > A\}$$

から，ただちに (ii) がでる．

(iii)　$\varepsilon > 0$ に対し，A を十分大きくとれば，

$$\sup_\lambda E(|X_\lambda|, |X_\lambda| > A) < \varepsilon.$$

このような A に対し $P(B) < \varepsilon/A$ とすれば，

$$\begin{aligned}E(|X_\lambda|, B) &\leqq E(|X_\lambda|, |X_\lambda| > A) + E(|X_\lambda|, \{|X_\lambda| \leqq A\} \cap B)\\ &< \varepsilon + AP(B) < 2\varepsilon\end{aligned}$$

が得られる．▌

この定理を用いて '$X_n \to X$ a.s.' と '$\|X_n - X\|_1 \to 0$' との関係を与えるつぎの定理を証明しよう．

定理 3.24　$X_1, X_2, \cdots, X \in L^1$ で $X_n \to X$ a.s. とする．このとき

$$\|X_n - X\|_1 \longrightarrow 0 \iff \{X_n\} \text{ が一様可積分}.$$

証明　$\|X_n - X\|_1 \to 0$ とすれば，

$$\|X_n\|_1 \leqq \|X\|_1 + \|X_n - X\|_1$$

により，$\|X_n\|_1$ が有界である．一般に $Y \in L^1$ に対しては

$$P(B) \longrightarrow 0 \Longrightarrow E(|Y|, B) \longrightarrow 0$$

であることは測度論でよく知られているから，$\varepsilon > 0$ に対して $\delta, \delta_1, \delta_2, \cdots$ が定まって

$$\begin{aligned}P(B) < \delta &\Longrightarrow E(|X|, B) < \varepsilon,\\ P(B) < \delta_n &\Longrightarrow E(|X_n|, B) < \varepsilon.\end{aligned}$$

さて

$$\begin{aligned}E(|X_n|, B) &\leqq E(|X|, B) + E(|X_n - X|, B)\\ &\leqq E(|X|, B) + \|X_n - X\|_1\end{aligned}$$

により，$\varepsilon > 0$ に対し n_0 を十分大きくとれば，すべての $n > n_0$ に対して $\|X_n - X\|_1 < \varepsilon$，したがって

$$P(B) < \delta \Longrightarrow E(|X_n|, B) < 2\varepsilon \qquad (n = n_0+1, n_0+2, \cdots).$$

したがって，

$$P(B) < \min(\delta, \delta_1, \cdots, \delta_{n_0}) \Longrightarrow \forall n \ E(|X_n|, B) < 2\varepsilon.$$

これで $\{X_n\}$ は前定理の (i), (iii) をみたすことがわかった．したがって上述の註により，$\{X_n\}$ は一様可積分である．

逆に $\{X_n\}$ が一様可積分であるとせよ．前定理の (ii) により，$A \to \infty$ のとき，n に関し一様に

$$P(\{|X_n| > A\} \cup \{|X| > A\}) \leqq P(|X_n| > A) + P(|X| > A) \longrightarrow 0$$

となる．左辺の集合を $B_{n,A}$ とおくと，$\sup_n P(B_{n,A}) \to 0$ $(n \to \infty)$．さて

$$\begin{aligned}\|X_n - X\|_1 &= E(|X_n - X|, B_{n,A}) + E(|X_n - X|, \{|X_n| \leqq A\} \cap \{|X| \leqq A\})\\ &\leqq E(|X_n|, B_{n,A}) + E(|X|, B_{n,A}) + E(|X_n - X| \wedge 2A).\end{aligned}$$

$P(B_{n,A}) \to 0$ であるから，前定理 (iii) により，第1項は A を大きくとれば一様に

小さくなる．$X \in L^1$ により，第2項も同様である．仮定 '$X_n \to X$ a.s.' と有界収束定理により，第3項も小さくなる．したがって $\|X_n - X\|_1$ は $n \to \infty$ のとき，0 に収束する．▮

実確率変数 X の確率法則 P^X は1次元分布であるから，$\mu = P^X$ に対し第2章で導入した分布関数 $F_\mu(x)$，能率 $M_p(\mu)$，絶対能率 $|M|_p(\mu)$，特性関数 $\varphi_\mu(z) \equiv \mathscr{F}\mu(z)$ が考えられるが，これをそれぞれ X の**分布関数**，**能率**，**絶対能率**，**特性関数**といい，$F_X(x)$, $M_p(X)$, $|M|_p(X)$, $\varphi_X(z)$ であらわす．積分の変換公式により，

$$\int_{R^1} f(x) P^X(dx) = E(f(X))$$

であるから，$f(x) = 1_{(-\infty, x]}$, x^p, $|x|^p$, e^{izx} とおいて

$$F_X(x) = P\{X \leqq x\}, \qquad M_p(X) = E(X^p),$$
$$|M|_p(x) = E(|X|^p), \qquad \varphi_X(z) = E(e^{izX})$$

が得られる．

$X \in L^1$ のときには

$$V(X) = E((X - EX)^2) = V(P^X)$$

を X の**分散**という．$X \in L^2$ のときには $V(X) < \infty$ となる．また

$$\sigma(X) = \sqrt{V(X)} = \sigma(P^X)$$

を X の**標準偏差**という．$X \in L^2$ のときには $\sigma(X) < \infty$ で，

Čebyšev の不等式　　$P\{|X - EX| \geqq a\sigma(X)\} \leqq \dfrac{1}{a^2}$

がなりたつ．これは

$$\sigma(X)^2 \geqq E((X - EX)^2, \ |X - EX| \geqq a\sigma(X))$$
$$\geqq a^2 \sigma(X)^2 P\{|X - EX| \geqq a\sigma(X)\}$$

から明らかである．

$X, Y \in L^2$ のとき

$$V(X, Y) = E((X - EX)(Y - EY))$$
$$= \iint_{R^2} (x - EX)(y - EY) P^{(X,Y)}(dxdy)$$

を X, Y の**共分散**といい，$\sigma(X), \sigma(Y)$ が正のとき

$$R(X, Y) = \frac{V(X, Y)}{\sigma(X)\sigma(Y)}$$

を X, Y の**相関係数**という．

平均値の加法性により，$V(X_i) < \infty$ $(i = 1, 2, \cdots, n)$ ならば

$$V\left(a_0 + \sum_{i=1}^n a_i X_i\right) = \sum_{i,j=1}^n a_i a_j V(X_i, X_j) \qquad (a_0, a_1, \cdots, a_n \text{ は実定数}).$$

定理 3.25　$X_1, X_2, \cdots, X_n$ が独立ならば ((i), (ii) では $X_i \in L^2$ を仮定)，

(i)　$V(X_i, X_j) = 0 \quad (i \neq j)$,

(ii)　$V\left(a_0 + \sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 V(X_i)$,

(iii)　$\varphi_{X_1 + X_2 + \cdots + X_n}(z) = \varphi_{X_1}(z) \varphi_{X_2}(z) \cdots \varphi_{X_n}(z)$,

(iv)　$P_{X_1 + X_2 + \cdots + X_n} = P_{X_1} * P_{X_2} * \cdots * P_{X_n}$.

証明　(i) と (iii) とは平均値の乗法性 (定理 3.15) により明らか．(ii) は (i) からすぐにでる．(iii) を用いて (iv) を示そう．

$$\mathscr{F}P_{X_1 + X_2 + \cdots + X_n}(z) = \varphi_{X_1 + X_2 + \cdots + X_n}(z) = \varphi_{X_1}(z) \varphi_{X_2}(z) \cdots \varphi_{X_n}(z)$$
$$= \prod_{i=1}^n \mathscr{F}P_{X_i}(z) = \mathscr{F}(P_{X_1} * P_{X_2} * \cdots * P_{X_n})(z).$$

対応 $\mu \leftrightarrow \mathscr{F}\mu$ の1対1性により (iv) が得られる．▮

確率ベクトル $X = (X_1, X_2, \cdots, X_n)$ に対しても

平均値ベクトル　　$EX = (EX_1, EX_2, \cdots, EX_n)$,

分散行列　　$V(X) = (V(X_i, X_j))_{i,j=1}{}^n$,

特性関数　　$\varphi(z) = E(e^{i(z,X)})$

$$\left(z = (z_1, z_2, \cdots, z_n), \quad (z, X) = \sum_{i=1}^n z_i X_i \ (\text{内積})\right)$$

を定義することができる．

定理 3.26 (Kac の定理)　$X_1, X_2, \cdots, X_n \in L^0$ が独立であるためには，すべての $z = (z_1, z_2, \cdots, z_n) \in \boldsymbol{R}^n$ に対し

$$E\left(\exp\left(i \sum_{k=1}^n z_k X_k\right)\right) = \prod_{i=1}^n E(e^{i z_k X_k})$$

がなりたつことが必要十分である．

証明　必要なことは平均値の乗法性から明らか．十分なことをいう．左辺は積分変換公式により $(\mathscr{F}P^X)(z)$ (X は $X_1, X_2, \cdots, X_n$ の結合変数) に等しい．右辺は Fubini の定理により

$$\prod_{i=1}^{n}\int_{R^1} e^{iz_kx_k}\mu_k(dx_k) \quad (\mu_k=P^{X_k})$$
$$=\int\int\cdots\int_{R^n} e^{i(z,x)}(\mu_1\times\mu_2\times\cdots\times\mu_n)(dx_1dx_2\cdots dx_n)$$
$$=\mathscr{F}(\mu_1\times\mu_2\times\cdots\times\mu_n)(z).$$

$\mu\leftrightarrow\mathscr{F}\mu$ の一意性 (n 次元のとき) により

$$P^X=\mu_1\times\mu_2\times\cdots\times\mu_n,$$

したがって

$$P(X_1\in E_1, X_2\in E_2, \cdots, X_n\in E_n)$$
$$=P^X(E_1\times E_2\times\cdots\times E_n)$$
$$=\mu_1(E_1)\mu_2(E_2)\cdots\mu_n(E_n)=\prod_{k=1}^{n}P(X_k\in E_k).$$

これは $X_1, X_2, \cdots, X_n$ の独立性を意味する. ▌

例題 3.7　(i)　$p>1$ のとき, $|M|_p(X_\lambda)$ $(\lambda\in\Lambda)$ が有界ならば, $\{X_\lambda, \lambda\in\Lambda\}$ は一様可積分であることを示せ.

［ヒント］　$|X_\lambda|\geqq A$ ならば $|X_\lambda|^p\geqq|X_\lambda|A^{p-1}$ となるから,

$$E(|X_\lambda|, |X_\lambda|\geqq A)\leqq E(|X_\lambda|^pA^{1-p}, |X_\lambda|\geqq A)\leqq A^{1-p}E(|X_\lambda|^p).$$

(ii)　$X_n\to X$ a.s., $X_n\in L^1$ $(n=1,2,\cdots)$ でしかも $X\notin L^1$ となる例をあげよ. また $X_n\to X$ a.s., $X_n\in L^1$ $(n=1,2,\cdots)$, $X\in L^1$ で, しかも $\|X_n-X\|_1\to 0$ とならない例をあげよ.

［ヒント］　$\Omega=(0,1)$, $P=$Lebesgue 測度とし

$$X_n(\omega)=\omega^{-(n-1)/n} \quad \text{および} \quad X_n(\omega)=n1_{(0,1/n)}(\omega)$$

を考察せよ.

(iii)　$\{X_\lambda, \lambda\in\Lambda\}$, $\{Y_\lambda, \lambda\in\Lambda\}$ がそれぞれ一様可積分であれば, $\{X_\lambda+Y_\lambda, \lambda\in\Lambda\}$ も一様可積分であることを示せ.

［ヒント］　定理 3.23 の註参照.

(iv)　**(Bienaymé の不等式)**　$f:[0,\infty)\to[0,\infty)$ が増加関数であるとき, すべての $a>0$ に対し

$$P(|X|\geqq a)\leqq\frac{1}{f(a)}E(f(|X|))$$

となることを示せ.

［ヒント］　$E(f(|X|))\geqq E(|f(X)|, |X|\geqq a)\geqq f(a)P(|X|\geqq a)$.

(v)　$X_1, X_2, \cdots$ を実確率変数の列とし, μ_n を X_n の確率法則とする.

$$X_n\longrightarrow 0 \text{ i.p.} \iff \mu_n\longrightarrow\delta.$$

［ヒント］　$P\{|X_n|>\varepsilon\}=\mu_n((-\varepsilon, \varepsilon)^c)$ と定理 2.10 (i)$\Leftrightarrow$(iv) とを用いて, '$\Longleftarrow$' がでる. '$\Longrightarrow$' は '確率収束$\Longrightarrow$法則収束' により明らか.

§3.8　条件付平均値作用素

(Ω, P) を基礎の確率空間とし, $\mathscr{B}$ を任意の σ 加法族 $\subset\mathscr{D}(P)$ としたとき, §3.5 において条件付確率測度 $P_{\mathscr{B}}=P_{\mathscr{B},\omega}$ を定義した.

定理 3.27　$\mathscr{B}_1$ と $\mathscr{B}_2$ とが同等ならば

$$P_{\mathscr{B}_1,\omega}=P_{\mathscr{B}_2,\omega} \text{ a.s.}$$

証明　$\mathscr{A}\subset\mathscr{D}(P)$ を Ω の上の可算分離族とする. 必要ならば, $\mathscr{A}$ にその中の任意の有限個の集合の交わりをつけ加えることにより, $\mathscr{A}$ は可算分離乗法族であるとしてもよい. $\mathscr{B}=\mathscr{B}_1\vee\mathscr{B}_2$ とおくと, $\mathscr{B}$ は $\mathscr{B}_1, \mathscr{B}_2$ のいずれとも同等である. $\mathscr{B}_1\subset\mathscr{B}$ であるから任意の $A\in\mathscr{D}(P)$ に対し

$$E(P_{\mathscr{B}_1}(A), B_1)=E(P_{\mathscr{B}}(A), B_1)\ (=P(A\cap B_1)), \qquad B_1\in\mathscr{B}_1$$

となる. 任意の $B\in\mathscr{B}$ に対して, これとほとんど一致する $B_1\in\mathscr{B}_1$ が存在するから

$$E(P_{\mathscr{B}_1}(A), B)=E(P_{\mathscr{B}}(A), B), \qquad B\in\mathscr{B}.$$

また §3.5 註 5 により

$$P_{\mathscr{B}_1}(A)\in\mathscr{B}_1\subset\mathscr{B} \text{ a.s.}, \qquad P_{\mathscr{B}}(A)\in\mathscr{B} \text{ a.s.}$$

であるから, 上の等式で B として

$$B=\{\omega\,|\,P_{\mathscr{B}_1,\omega}(A)>P_{\mathscr{B},\omega}(A)\}, \quad \{\omega\,|\,P_{\mathscr{B},\omega}(A)>P_{\mathscr{B}_1,\omega}(A)\}$$

とおいて,

$$P_{\mathscr{B}_1,\omega}(A)=P_{\mathscr{B},\omega}(A) \text{ a.s.}$$

を証明することができる. 同様に

$$P_{\mathscr{B}_2,\omega}(A)=P_{\mathscr{B},\omega}(A) \text{ a.s.},$$

したがって

$$P_{\mathscr{B}_1,\omega}(A)=P_{\mathscr{B}_2,\omega}(A) \text{ a.s.}$$

$\mathscr{A}$ が可算族であるから, 例外の P 零集合 N は $A\in\mathscr{A}$ に無関係にとれる. すな

わち $\omega \in \Omega - N$ に対して

$$P_{\mathcal{B}_1,\omega}(A) = P_{\mathcal{B}_2,\omega}(A), \qquad A \in \mathcal{A}.$$

確率測度一致の定理により，$\omega \in \Omega - N$ に対し $P_{\mathcal{B}_1,\omega}, P_{\mathcal{B}_2,\omega}$ は $\sigma[\mathcal{A}]$ の上で一致し，定理 3.1* により，完全に一致する．▌

さて個々の ω に対し $(\Omega, P_{\mathcal{B},\omega})$ は可分完全確率空間であるから，この上で事象，確率変数，確率法則（**条件付確率法則**という），平均値（**条件付平均値**という）などの確率論的概念を定義することができる．$Y = Y(\omega')$ を (Ω, P) の上の S 値確率変数とするとき，ほとんどすべての ω に対し，$Y(\omega')$ は $(\Omega, P_{\mathcal{B},\omega})$ の上で S 値確率変数である．（$Y(\omega')$ で ω のかわりに ω' を用いたのは，$P_{\mathcal{B},\omega}$ の ω との混同をさけるためで，これは

$$\int_0^t t dt = \frac{t^2}{2}$$

とかくよりも

$$\int_0^t s ds = \frac{t^2}{2}$$

とかく方がよいというのと同じ理由によるのである．）$Y(\omega')$ は (Ω, P) の上の S 値確率変数であるから，S の上の可算分離族 $\{E_n\}$ があって，

$$Y^{-1}(E_n) \in \mathcal{D}(P), \qquad n = 1, 2, \cdots.$$

これから Ω の P 零集合 N があって，$\omega \in \Omega - N$ に対し

$$Y^{-1}(E_n) \in \mathcal{D}(P_{\mathcal{B},\omega}), \qquad n = 1, 2, \cdots$$

が得られ，$Y(\omega')$ が $(\Omega, P_{\mathcal{B},\omega})$ の上の S 値確率変数となる．

$Y(\omega')$ が (Ω, P) の上の拡張された実確率変数のときには，ほとんどすべての ω に対し $Y(\omega')$ は $(\Omega, P_{\mathcal{B},\omega})$ の上でも同様となり，

$$E_{\mathcal{B}}(Y) = E_{\mathcal{B},\omega}(Y) = \int_\Omega Y(\omega') P_{\mathcal{B},\omega}(d\omega')$$

が考えられる．$E(Y)$ と同様に $E_{\mathcal{B}}(Y) = E_{\mathcal{B},\omega}(Y)$ も常に確定するわけではない．しかし $Y(\omega') \geqq 0$ のときにはともに確定し

$$E(E_{\mathcal{B}}(Y)) = E(Y)$$

となることが証明される．実際 $Y = 1_A$ $(A \in \mathcal{D}(P))$ のときには

$$E_{\mathcal{B},\omega}(Y) = P_{\mathcal{B},\omega}(A), \qquad E(Y) = P(A)$$

となるから，$P_{\mathcal{B}}$ の定義により $E(E_{\mathcal{B}}(Y)) = E(P_{\mathcal{B}}(A)) = P(A)$．これからいつもの論法により $Y \geqq 0$ のときには上の等式を得る．もし $E(Y) < \infty$ ならば，上の等式から

$$E_{\mathcal{B}}(Y) < \infty \text{ a.s.}$$

がでるから，$E(Y)$ が確定するとき（すなわち EY^+, EY^- の少なくとも一方が有限のとき）にはほとんどすべての ω に対し $E_{\mathcal{B}}Y$ が確定する（$E_{\mathcal{B}}Y$ が確定しない ω に対しては $E_{\mathcal{B}}Y$ は任意に定めておく）．さらに

$$|E(Y)| < \infty \Leftrightarrow E|Y| < \infty$$
$$\Rightarrow E_{\mathcal{B}}|Y| < \infty \text{ a.s.} \Leftrightarrow |E_{\mathcal{B}}Y| < \infty \text{ a.s.}$$

も得られる．しかし $E_{\mathcal{B}}|Y|$ の有限性から必ずしも $E|Y|$ の有限性がでないことに注意すべきである．

$E(Y)$ が有限のとき，$E_{\mathcal{B}}Y$ はほとんど確実に有限であるが，例外の P 零集合は Y に関係する．しかし可算個の $Y_1, Y_2, \cdots$ に対しては例外集合を共通にとることができる．

$E(Y, A) = E(Y1_A)$，$E_{\mathcal{B}}(Y, A) = E_{\mathcal{B}}(Y1_A)$ により，$E_{\mathcal{B}}(Y, A)$ についても上と同様のことがいえる．

定理 3.28 平均値の性質（定理 3.5）は条件付平均値 $E_{\mathcal{B}}$ についてほとんど確実になりたつ．例えば無限和との交換定理 (viii) の後半はつぎのようになる．

$$E\left(\sum_n |X_n|\right) < \infty \Longrightarrow E_{\mathcal{B}}\left(\sum_n |X_n|\right) < \infty \text{ a.s.}$$
$$\Longrightarrow E_{\mathcal{B}}\left(\sum_n X_n\right) = \sum_n E_{\mathcal{B}} X_n \text{ a.s.} \qquad \text{——}$$

L^1 の元で，$\mathcal{B}$ に関して可測な確率変数とほとんど一致するものの全体を $L_{\mathcal{B}}{}^1$ であらわす．L^1 の中ではほとんど一致する確率変数は同一視されるから，$L_{\mathcal{B}}{}^1 = L_{\mathcal{B} \vee \mathfrak{z}}{}^1$ である．したがって $\mathcal{B}_1, \mathcal{B}_2$ が同等なときには $L_{\mathcal{B}_1}{}^1 = L_{\mathcal{B}_2}{}^1$ となる．

上述の定義では $E_{\mathcal{B}}X$ は $P_{\mathcal{B}}$ から導かれたが，$\tilde{X} = E_{\mathcal{B}}X$ はつぎの 2 条件で特長づけられるから，$X \in L^1$ に対しては，これで $E_{\mathcal{B}}X$ を定義してもよい．

(CE. 1) $\qquad \tilde{X} \in L_{\mathcal{B}}{}^1$,

(CE. 2) $\qquad E(\tilde{X}, B) = E(X, B), \qquad B \in \mathcal{B}$.

実際 $X = 1_A$ のときには，$E_{\mathcal{B}}X = P_{\mathcal{B}}(A)$ となるから，$\tilde{X} = E_{\mathcal{B}}X$ がこの 2 条件をみ

たすことは，$P_{\mathcal{B}}$ の定義から明らかである．これから1次結合，極限をとることによって，一般の $X\in L^1$ に対して $\tilde{X}=E_{\mathcal{B}}X$ が上の条件をみたす．逆に上の2条件をみたす任意の $\tilde{X}$ に対して

$$E(\tilde{X}, B) = E(X, B) = E(E_{\mathcal{B}}X, B), \qquad B\in\mathcal{B}$$

となる．$\tilde{X}, E_{\mathcal{B}}X\in L_{\mathcal{B}}{}^1$ であるから，

$$B_+ = \{\tilde{X}(\omega) > E_{\mathcal{B},\omega}X\}, \qquad B_- = \{\tilde{X}(\omega) < E_{\mathcal{B},\omega}X\}$$

は $\mathcal{B}$ に属する．上の等式で $B=B_+$ とおくと

$$E(\tilde{X}, B_+) = E(E_{\mathcal{B}}X, B_+).$$

B_+ の上では $\tilde{X}>E_{\mathcal{B}}X$ であるから，上の式から $P(B_+)=0$ がでる．同様に $P(B_-)=0$．したがって $\tilde{X}$ と $E_{\mathcal{B}}X$ とはほとんど一致する．これで $\tilde{X}=E_{\mathcal{B}}X$ が上の2条件で特長づけられることがわかった．

前節では $L^p=L^p(\Omega, P)$ を $1\leqq p<\infty$ に対して導入したが，$L^\infty=L^\infty(\Omega, P)$ を

$$L^\infty = \{X\in L^0 \mid \|X\|_\infty<\infty\}$$

で定義する．ここで

$$\|X\|_\infty = \operatorname*{ess.sup}_{\omega\in\Omega}|X(\omega)| = \inf\{b \mid |X(\omega)|\leqq b \text{ a.s.}\}$$

$$(\text{ess.sup}=\text{essential supremum})$$

である．L^∞ はノルム $\|\ \|_\infty$ に関して Banach 空間である．L^p は可分であったが L^∞ は一般に可分でない．Ω が可算集合のときには L^∞ は可分であるが，$\Omega=[0, 1)$，$P=$Lebesgue 測度，のときには，$[\alpha, \beta)\neq[\gamma, \delta)$ のとき

$$\|1_{[\alpha,\beta)}-1_{[\gamma,\delta)}\|_\infty = 1$$

であるから，L^∞ は可分でない．

前節で $\|X\|_p$ は $1\leqq p<\infty$ で単調増大であることをのべたが，$1\leqq p\leqq\infty$ でもそうであることは明らかである．したがって，L^p は $1\leqq p\leqq\infty$ で単調減少である．

$L_{\mathcal{B}}{}^1$ と同様に $L_{\mathcal{B}}{}^p$ を

$$L_{\mathcal{B}}{}^p = \{X\in L^p \mid X\in\mathcal{B} \text{ a.s.}\}$$

で定義する．明らかに

$$L_{\mathcal{B}}{}^p = L_{\mathcal{B}}{}^1\cap L^p$$

である．

$p, q\geqq 1$ が共役であるとは

$$\frac{1}{p}+\frac{1}{q} = 1 \qquad \left(\text{ただし } \frac{1}{\infty}=0 \text{ とする}\right)$$

がなりたつことである．$1\leqq p<\infty$ のときには補題2.4と同じ方法で，Hölder の不等式

$$E(|X\cdot Y|) \leqq \|X\|_p\|Y\|_q$$

が証明される．とくに $\|X\|_p, \|Y\|_q<\infty$ のときには，$X\cdot Y\in L^1$ で

$$\|X\cdot Y\|_1 \leqq \|X\|_p\|Y\|_q$$

となる．これらの事実は $p=\infty$（したがって $q=1$）のときにもなりたつことはすぐにわかる．

定理 3.29 (Jensen の不等式) $X\in L^1(\Omega, P)$ の像集合 $X(\Omega)$ が区間 I（開，閉，半開いずれでもよい）に含まれるとする．f を I の上の凸関数とすれば

$$f(EX) \leqq E(f(X)),$$
$$f(E_{\mathcal{B}}X) \leqq E_{\mathcal{B}}(f(X)) \text{ a.s.}$$

（右辺は ∞ になり得る．）

証明 $X(\Omega)\subset I$ により，$m\equiv EX\in I$．f は凸であるから，$y=f(x)$ のグラフは，その上の点 $(m, f(m))$ を通る適当な直線

$$y = f(m)+\alpha(x-m)$$

の上にある．したがって

$$f(X(\omega)) \geqq f(m)+\alpha(X(\omega)-m), \qquad \omega\in\Omega.$$

$X\in L^1$ であるから，右辺も L^1 に属する．ゆえに

$$E\{(f(X))^-\} \leqq E(|f(m)+\alpha(X-m)|) < \infty.$$

これは $E(f(X))$ が確定し，$(-\infty, \infty]$ の中の値をとることを意味する．$f(X)\geqq f(m)+\alpha(X-m)$ の両辺の平均値をとって

$$E(f(X)) \geqq f(m) = f(EX).$$

これで第1式は証明された．

第2式をいうには $X\in L^1(\Omega, P)$ により，$|X|\in L^1$．(CE. 2) で X, B のかわりに $|X|, \Omega$ とおいて

$$E(E_{\mathcal{B}}(|X|)) = E|X| < \infty.$$

したがって

$$E_{\mathcal{B}}|X| < \infty \text{ a.s.} \quad \text{すなわち} \quad X\in L^1(\Omega, P_{\mathcal{B}}) \text{ a.s.}$$

さて上に証明した第1式で P のかわりに $P_{\mathcal{B}}$ とおいて，第2式を得る．▌

この定理で $I=(-\infty,\infty)$，$f(x)=|x|^p$ $(1\leqq p<\infty)$ とおいて，$X\in L^p$ $(\subset L^1)$ のとき

$$|E_{\mathcal{B}}X|^p\leqq E_{\mathcal{B}}|X|^p,$$
$$E(|E_{\mathcal{B}}X|^p)\leqq E(E_{\mathcal{B}}|X|^p)=E(|X|^p)<\infty$$

を得る．これから

$$X\in L^p\Longrightarrow E_{\mathcal{B}}X\in L^p\Longrightarrow E_{\mathcal{B}}X\in L_{\mathcal{B}}{}^p,\qquad \|E_{\mathcal{B}}X\|_p\leqq\|X\|_p$$

が得られる．($p=\infty$ のときにはこの事実は明らかである.) これは $1\leqq p\leqq\infty$ に対し

$$E_{\mathcal{B}}\colon L^p\longrightarrow L^p,\qquad E_{\mathcal{B}}(L^p)\subset L_{\mathcal{B}}{}^p$$

を意味する．しかも $X\in L_{\mathcal{B}}{}^p\subset L_{\mathcal{B}}{}^1$ のときには，$\tilde{X}=X$ は (CE. 1), (CE. 2) をみたすことは明らかであるから，$E_{\mathcal{B}}X=X$ となる．すなわち

$$X\in L_{\mathcal{B}}{}^p\Longrightarrow E_{\mathcal{B}}X=X.$$

これから

$$E_{\mathcal{B}}(L^p)=L_{\mathcal{B}}{}^p$$

が得られる．

定理 3.30 p,q を共役とし，$X\in L^p$ とする．$\tilde{X}=E_{\mathcal{B}}X$ となるための必要十分条件は

(i) $\tilde{X}\in L_{\mathcal{B}}{}^p$,

(ii) $E(\tilde{X}Y)=E(XY)$, $\quad Y\in L_{\mathcal{B}}{}^q$

がなりたつことである．

証明 $B\in\mathcal{B}$ のときには $1_B\in L_{\mathcal{B}}{}^q$ であるから，十分性は明らか．$\tilde{X}=E_{\mathcal{B}}X$ が (i), (ii) をみたすことをいえば，必要性がいえる．(i) は明らか．$Y=1_B$ $(B\in\mathcal{B})$ のときには (ii) は (CE. 2) から明らか．一般の $Y\in L_{\mathcal{B}}{}^q$ はこのような関数の1次結合の列の極限 ($\|\ \|_q$ による) であらわされることに注意し，Hölder の不等式を用いると，一般の $Y\in L_{\mathcal{B}}{}^q$ に対して $\tilde{X}=E_{\mathcal{B}}X$ が (ii) をみたすことがいえる．▌

定理 3.31 p,q を共役とし，$X\in L^p$，$Y\in L^q$ とする．

(i) $E((E_{\mathcal{B}}X)Y)=E(X(E_{\mathcal{B}}Y))=E((E_{\mathcal{B}}X)(E_{\mathcal{B}}Y))$.

(ii) $Y\in L_{\mathcal{B}}{}^q\Longrightarrow E_{\mathcal{B}}(XY)=YE_{\mathcal{B}}X$.

証明 (i) $E_{\mathcal{B}}Y\in L_{\mathcal{B}}{}^q$ であるから，前定理により

$$E(XE_{\mathcal{B}}Y)=E((E_{\mathcal{B}}X)(E_{\mathcal{B}}Y)).$$

p と q，X と Y とを交換して考えると

$$E(YE_{\mathcal{B}}X)=E((E_{\mathcal{B}}Y)(E_{\mathcal{B}}X)).$$

これで (i) が証明された．

(ii) $X\in L^p$，$Y\in L_{\mathcal{B}}{}^q\subset L^q$ により $XY\in L^1$，$YE_{\mathcal{B}}X\in L_{\mathcal{B}}{}^1$．任意の $Z\in L_{\mathcal{B}}{}^\infty$ に対し，$YZ\in L_{\mathcal{B}}{}^q$，前定理 (必要性) により，

$$E(XYZ)=E((E_{\mathcal{B}}X)YZ).$$

前定理 (十分性) により

$$E_{\mathcal{B}}(XY)=(E_{\mathcal{B}}X)Y.\qquad ▌$$

定理 3.32 $X\in L^2$ に対し $\tilde{X}=E_{\mathcal{B}}X$ は $\|X-\tilde{X}\|_2$ を最小にする $L_{\mathcal{B}}{}^2$ の元として特長づけられる．したがって $E_{\mathcal{B}}\colon L^2\to L^2$ は Hilbert 空間 L^2 からその閉部分空間 $L_{\mathcal{B}}{}^2$ への射影作用素と一致する．

証明 任意の $\tilde{X}\in L_{\mathcal{B}}{}^2$ に対し

$$\|X-\tilde{X}\|_2{}^2=\|(X-E_{\mathcal{B}}X)-(\tilde{X}-E_{\mathcal{B}}X)\|_2{}^2$$
$$=\|X-E_{\mathcal{B}}X\|_2{}^2+2E((X-E_{\mathcal{B}}X)(\tilde{X}-E_{\mathcal{B}}X))+\|\tilde{X}-E_{\mathcal{B}}X\|_2{}^2,$$
$$X-E_{\mathcal{B}}X\in L^2,\qquad \tilde{X}-E_{\mathcal{B}}X\in L_{\mathcal{B}}{}^2,\qquad E_{\mathcal{B}}(X-E_{\mathcal{B}}X)=0$$

により，定理 3.30 (必要性) を用いて

$$E((X-E_{\mathcal{B}}X)(\tilde{X}-E_{\mathcal{B}}X))=E((E_{\mathcal{B}}(X-E_{\mathcal{B}}X))(\tilde{X}-E_{\mathcal{B}}X))=0$$

を得る．したがって，

$$\|X-\tilde{X}\|_2{}^2=\|X-E_{\mathcal{B}}X\|_2{}^2+\|\tilde{X}-E_{\mathcal{B}}X\|_2{}^2\qquad(\tilde{X}\in L_{\mathcal{B}}{}^2)$$
$$\geqq\|X-E_{\mathcal{B}}X\|_2{}^2$$

で，等号は $\tilde{X}=E_{\mathcal{B}}X$ $(\in L_{\mathcal{B}}{}^2)$ のときに成立する．▌

以上の議論では $\mathcal{B}$ を固定していたが，$\mathcal{B}$ が変化するとき，$E_{\mathcal{B}}$ が受ける影響を調べて見よう．

定理 3.33 $\mathcal{B}_1\subset\mathcal{B}_2$ のとき

$$E_{\mathcal{B}_1}E_{\mathcal{B}_2}=E_{\mathcal{B}_1}.$$

証明 $X\in L^p$ のときには $E_{\mathcal{B}_2}X\in L_{\mathcal{B}_2}{}^p\subset L^p$，したがって

$$E_{\mathcal{B}_1}(E_{\mathcal{B}_2}X)\in L_{\mathcal{B}_1}{}^p.$$

p に共役な q をとり，Y を $L_{\mathcal{B}_1}{}^q$ の任意の元とする．$\mathcal{B}_1\subset\mathcal{B}_2$ により

$$Y\in L_{\mathcal{B}_1}{}^q\subset L_{\mathcal{B}_2}{}^q.$$

定理 3.30 (必要性) により

$$E((E_{\mathcal{B}_1}(E_{\mathcal{B}_2}X))\,Y) = E((E_{\mathcal{B}_2}X)\,Y) = E(XY).$$

再び上の定理 (十分性) により

$$E_{\mathcal{B}_1}(E_{\mathcal{B}_2}X) = E_{\mathcal{B}_1}X. \qquad \blacksquare$$

U を確率変数としたとき，$\mathcal{B}=\bar{\sigma}[U]$ に対応する $E_{\mathcal{B}}X$ を E_UX または $E(X|U)$ であらわす．U が位相確率変数のときには $\bar{\sigma}[U]$ は $\sigma[U]$ と同等であるから，$\mathcal{B}=\sigma[U]$ に対応する $E_{\mathcal{B}}X$ を E_UX または $E(X|U)$ であらわすといってもよい．上の定理を適用して

$$E_UE_{(U,V)} = E_U.$$

$\mathcal{B}=\bar{\sigma}[U]$ のときには

$$L_{\mathcal{B}}{}^p = \{X\in L^p \mid X \text{ は } U \text{ の娘}\}$$

となるから，$X\in L^2$ のときには

$$\|X-\varphi(U)\|_2 = \text{最小}, \qquad \varphi\in L^2(S, P^U) \quad (S \text{ は } U \text{ の値域})$$

にする $\varphi(U)$ が $E(X|U)$ を与える．

例題 3.8　(i)　$X\in L^p$ が $\mathcal{B}$ と独立 (すなわち $\sigma[X]$ が $\mathcal{B}$ と独立) のとき

$$E_{\mathcal{B}}X = EX$$

を証明せよ．

(ii)　$\mathcal{B}_1\supset\mathcal{B}_2\supset\mathcal{B}_3\supset\cdots$，$\mathcal{B}=\bigcap_n\mathcal{B}_n$ のとき，$X\in L^2$ に対し

$$\|E_{\mathcal{B}}X-E_{\mathcal{B}_n}X\|_2{}^2 \longrightarrow 0$$

を証明せよ．

［ヒント］　$L_{\mathcal{B}_1}{}^2\supset L_{\mathcal{B}_2}{}^2\supset\cdots$，$L_{\mathcal{B}}{}^2=\bigcap_n L_{\mathcal{B}_n}{}^2$ に注意し，定理 3.32 を用いると，Hilbert 空間 L^2 の上の問題に帰着される．

(iii)　(X, Y) の確率法則が 2 次元の Gauss 分布

$$N(m, v), \qquad m = (m_1, m_2), \quad v = (v_{ij})_{i,j=1}{}^2$$

のとき，E_XY を求めよ．

［ヒント］　まず $P_{X=x}(Y\in E)$ を求め，つぎに積分して $E_{X=x}(Y)$ を求め，ここで x のかわりに $X=X(\omega)$ をおきかえると E_XY が得られる．$\det v>0$ のときには $N(m, v)$ が連続な密度をもつから，$P_{X=x}(Y\in E)$ が例題 3.5 (ii) の考えを用いて得られる．$\det v=0$ のときには $N(m, v)$ がある直線の上に集中していることに注意せよ．

第4章　独立確率変数の和

独立な実確率変数の和の理論は大数の法則の一般化，精密化のために開発されたもので，1930 年代には確率論で最も重要な分野であった．この理論のために多数の有力な手段が A. Kolmogorov, A. Hinčin, P. Lévy などによって導入され，現代確率論の土台となっている．

§4.1　一般的事項

(Ω, P) を基礎の可分完全確率測度空間とし，$X_n(\omega)$ $(n=1, 2, \cdots)$ を (Ω, P) の上の独立な実確率変数の (無限) 列とする．X_n の確率法則を μ_n とする．$X_n(\omega)$ $(n=1, 2, \cdots)$ の結合変数

$$X(\omega) = (X_1(\omega), X_2(\omega), \cdots)$$

は無限次元の確率ベクトルと考えられ，その確率法則は μ_n $(n=1, 2, \cdots)$ の完備直積：

$$\mu = \overline{\mu_1\times\mu_2\times\cdots}$$

である (定理 3.14 (iii))．μ は位相確率変数 X の確率法則であるから，位相空間 $\boldsymbol{R}^\infty$ の上の可分完全正則確率測度である (§3.2 参照)．

逆に $\mu_1, \mu_2, \cdots$ が任意に与えられた 1 次元分布の列とすると，適当な可分完全確率空間 (Ω, P) の上に独立な実確率変数列 $X_n(\omega)$ $(n=1, 2, \cdots)$ を適当にとって，おのおのの X_n の確率法則が μ_n に等しいようにできる．この方法はいろいろあるが，

$$\Omega = \boldsymbol{R}^\infty, \qquad P = \overline{\mu_1\times\mu_2\times\cdots},$$

$$X_n(\omega) = p_n(\omega) \quad (=\omega \text{ の第 } n \text{ 座標})$$

もその一つであって，これを**座標表現**という．

$X_1, X_2, \cdots$ が独立な実確率変数の列とするとき，その部分和の列

$$S_n = X_1+X_2+\cdots+X_n, \qquad n = 1, 2, \cdots$$

がどういう行動をするかを調べるのが本章の目的である．

S_n が収束するという事象の外延 C は

$$C = \left\{(\forall m)(\exists n)(\forall j,\, k>n)|S_j(\omega)-S_k(\omega)|<\frac{1}{m}\right\}$$

$$= \bigcap_m \bigcup_n \bigcap_{j>k>n} \left\{|S_j(\omega)-S_k(\omega)|<\frac{1}{m}\right\}$$

であり，$S_r(\omega)$ は ω の P 可測関数であるから，$|S_j(\omega)-S_k(\omega)|$ も同様である．したがって C は P 可測集合であって，$P(C)$ が考えられる．これが S_n が収束する確率をあらわしている．

Kolmogorov の 0-1 法則を用いて

$$P(C) = 0 \text{ または } 1$$

を証明しよう．実確率変数 Y が σ 加法族 $\mathcal{B}$ $(\subset \mathcal{D}(P))$ に関して可測なこと，すなわち $Y \in \mathcal{B}/\mathcal{B}^1$ を簡単に $Y \in \mathcal{B}$ とかくことにする．明らかに

$$Y_1, Y_2, \cdots, Y_n \in \mathcal{B} \Longrightarrow Y_1+Y_2+\cdots+Y_n \in \mathcal{B}$$

である．さて上に与えられた $X_1, X_2, \cdots$ に対し $\sigma[X_1], \sigma[X_2], \cdots$ を $\mathcal{B}_1, \mathcal{B}_2, \cdots$ であらわそう．$X_n \in \mathcal{B}_n$ であるから

$$X_n \in \mathcal{B}_{k\infty} \equiv \bigvee_{j>k} \mathcal{B}_j \qquad (n>k)$$

がなりたつ．したがって

$$S_j-S_k = X_{k+1}+X_{k+2}+\cdots+X_j \in \mathcal{B}_{k\infty} \qquad (j>k),$$

$$\left\{|S_j-S_k|<\frac{1}{m}\right\} \in \mathcal{B}_{k\infty} \qquad (j>k),$$

$$\bigcap_{j>k>n} \left\{|S_j-S_k|<\frac{1}{m}\right\} \in \mathcal{B}_{n\infty}.$$

さて "$S_1, S_2, \cdots$ が収束する" のは "$S_{n+1}, S_{n+2}, \cdots$ が収束する" のと同等であるから

$$C = \bigcap_m \bigcup_{r>n} \bigcap_{j>k>r} \left\{|S_j-S_k|<\frac{1}{m}\right\}.$$

上に注意したことから，すべての n に対し

$$C \in \mathcal{B}_{n\infty},$$

したがって

$$C \in \bigwedge_n \mathcal{B}_{n\infty} = \bigwedge_n \bigvee_{k>n} \mathcal{B}_k.$$

$X_1, X_2, \cdots$ が独立であるから，$\mathcal{B}_1, \mathcal{B}_2, \cdots$ が独立，したがって Kolmogorov の 0-1

法則により，

$$P(C) = 0 \text{ または } 1$$

が得られる．この事実は独立な確率変数を項とする級数はほとんど確実に収束する（**概収束**）かまたはほとんど確実に発散する（**概発散**）かのいずれかであって，確率 1/2 で収束し，確率 1/2 で発散するという中間の場合は絶対におこらないという著しい事実をあらわしている．

ではいかなる条件のもとで概収束となるか，また概発散のときにその発散の状況はどの程度のものかという問題がおこる．以下の諸節でこれらの問題に答えよう．

$X_1, X_2, \cdots$ が独立で，その確率法則がそれぞれ $\mu_1, \mu_2, \cdots$ とすると，$S_n=X_1+X_2+\cdots+X_n$ の確率法則は $\mu_1*\mu_2*\cdots*\mu_n$ であり，S_n の特性関数は $\mathcal{F}\mu_1(z)\mathcal{F}\mu_2(z)\cdots\mathcal{F}\mu_n(z)$ である．したがって S_n $(n=1, 2, \cdots)$ の行動をしらべるのに

$$\nu_n = \mu_1*\mu_2*\cdots*\mu_n, \qquad \mathcal{F}\nu_n(z) = \mathcal{F}\mu_1(z)\mathcal{F}\mu_2(z)\cdots\mathcal{F}\mu_n(z), \qquad n = 1, 2, \cdots$$

の行動をしらべることが重要な役割をすることがわかる．

後に利用するため独立確率変数の和に関する Kolmogorov の不等式と Ottaviani の不等式を証明しておく．

定理 4.1 (**Kolmogorov の不等式**) $X_1, X_2, \cdots, X_n$ が独立で

$$EX_i = 0, \qquad v_i \equiv V(X_i) < \infty$$

とするとき，

$$P\left\{\max_{k=1}^{n} |X_1+X_2+\cdots+X_k| \geqq a\right\} \leqq \frac{1}{a^2}\sum_{k=1}^{n} v_k \qquad (a>0).$$

証明 $S_k=X_1+X_2+\cdots+X_k$ とおき，$T_k=\max_{i=1}^{k} |S_i|$ とおく．

$$A_k = \{T_{k-1}<a,\, |S_k|\geqq a\} \qquad (T_0=0)$$

とおくと，$A_1, A_2, \cdots, A_n$ は互いに素で，

$$\{T_n \geqq a\} = A_1+A_2+\cdots+A_n.$$

A_k の指示関数を $e_k(\omega)$ とし，$\mathcal{B}_k=\sigma[X_k]$ とおく．$X_1, X_2, \cdots, X_n$ の独立性により

$$V(S_n) = \sum_{i=1}^{n} v_i.$$

したがって証明すべき不等式は

$$\sum_{k=1}^{n} P(A_k) \leqq \frac{1}{a^2} V(S_n).$$

$A_1, A_2, \cdots, A_n$ が互いに素であるから $\sum_{k=1}^{n} e_k \leqq 1$. 仮定 $EX_k=0$ により, $ES_n=0$. ゆえに

$$V(S_n) = E(S_n^2) \geqq \sum_{k=1}^{n} E(S_n^2 e_k),$$

$$E(S_n^2 e_k) = E(S_k^2 e_k) + 2E(S_k(S_n-S_k)e_k) + E((S_n-S_k)^2 e_k).$$

$\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_n$ は独立で

$$S_k e_k \in \mathcal{B}_1 \vee \mathcal{B}_2 \vee \cdots \vee \mathcal{B}_k,$$

$$S_n - S_k = X_{k+1} + X_{k+2} + \cdots + X_n \in \mathcal{B}_{k+1} \vee \mathcal{B}_{k+2} \vee \cdots \vee \mathcal{B}_n$$

であるから, 定理 3.12 により $S_k e_k$, $S_n - S_k$ が独立. $E(S_n - S_k) = 0$ であるから,

$$E(S_k(S_n-S_k)e_k) = E(S_k e_k) E(S_n - S_k) = 0.$$

これを上の等式に入れて

$$E(S_n^2 e_k) \geqq E(S_k^2 e_k) = E(S_k^2, A_k),$$

A_k の上では $|S_k| \geqq a$ であるから

$$E(S_n^2 e_k) \geqq a^2 P(A_k).$$

したがって

$$V(S_n) \geqq a^2 \sum_{k=1}^{n} P(A_k).$$ ∎

定理 4.2 (**Ottaviani の不等式**)　$X_1, X_2, \cdots, X_n$ が独立で

$$P\{|X_{k+1} + X_{k+2} + \cdots + X_n| \leqq a\} \geqq \beta, \qquad k = 0, 1, 2, \cdots, n-1$$

ならば,

$$P\left\{\max_{k=1}^{n} |X_1 + X_2 + \cdots + X_k| > 2a\right\} \leqq \frac{1}{\beta} P\{|X_1 + X_2 + \cdots + X_n| > a\}.$$

ここで a, β は正の定数である.

証明　$S_k, T_k, \mathcal{B}_k$ は前定理の証明と同様とし,

$$A_k = \{T_{k-1} \leqq 2a, |S_k| > 2a\},$$

$$B_k = \{|S_n - S_k| \leqq a\} \qquad (\text{ただし } B_n = \Omega)$$

とおくと, $A_1, A_2, \cdots, A_n$ は素で

$$\{T_n > 2a\} = A_1 + A_2 + \cdots + A_n.$$

したがって証明すべき不等式は

$$\sum_{k=1}^{n} P(A_k) \leqq \frac{1}{\beta} P\{|S_n| > a\}$$

である.

$A_1 \cap B_1$, $A_2 \cap B_2$, $\cdots$, $A_n \cap B_n$ は互いに素で, しかも

$$\omega \in A_k \cap B_k \Longrightarrow |S_k(\omega)| > 2a, \ |S_n(\omega) - S_k(\omega)| \leqq a$$

$$\Longrightarrow |S_n(\omega)| > a$$

であるから

$$\sum_{k=1}^{n} A_k \cap B_k \subset \{|S_n| > a\}.$$

前定理の証明と同様に,

$$S_k, T_k \in \mathcal{B}_1 \vee \mathcal{B}_2 \vee \cdots \vee \mathcal{B}_k, \qquad S_n - S_k \in \mathcal{B}_{k+1} \vee \mathcal{B}_{k+2} \vee \cdots \vee \mathcal{B}_n,$$

したがって

$$A_k \in \mathcal{B}_1 \vee \mathcal{B}_2 \vee \cdots \vee \mathcal{B}_k, \qquad B_k \in \mathcal{B}_{k+1} \vee \mathcal{B}_{k+2} \vee \cdots \vee \mathcal{B}_n$$

となることから, 定理 3.12 により

$$P(A_k \cap B_k) = P(A_k) P(B_k) \geqq P(A_k)\beta$$

がでる. これを念頭において

$$P\{|S_n| > a\} \geqq \sum_{k=1}^{n} P(A_k \cap B_k) \geqq \beta \sum_{k=1}^{n} P(A_k) = \beta P\{T_n > 2a\}.$$ ∎

例題 4.1　(i)　μ, ν が1次元分布で $M_2(\mu), M_2(\nu) < \infty$ のとき,

$$M(\mu * \nu) = M(\mu) + M(\nu), \qquad V(\mu * \nu) = V(\mu) + V(\nu).$$

[ヒント]　例えば座標表現によって独立な確率変数 X, Y を $P^X = \mu$, $P^Y = \nu$ となるように定めると, $P^{X+Y} = \mu * \nu$.

したがって

$$M(\mu * \nu) = E(X+Y), \qquad V(\mu * \nu) = V(X+Y).$$

(ii)　Ottaviani の不等式は $X_1, X_2, \cdots, X_n$ が確率ベクトル変数のときにもなりたつことを証明せよ.

§4.2 独立確率変数の級数の概収束

$X_1, X_2, \cdots$ を独立実確率変数とするとき，級数

$$\sum_n X_n$$

は概収束するか概発散するかのいずれかであることを前節でのべた．本節では概収束するための条件を考察する．

定理 4.3 (Kolmogorov の定理) $\sum_n E(X_n)$，$\sum_n V(X_n)$ がともに収束すれば，$\sum_n X_n$ は概収束する．

証明 $X_n=(X_n-E(X_n))+E(X_n)$，$V(X_n)=V(X_n-E(X_n))$ であるから，$EX_n=0$ として一般性を失わない．

$$S_n = X_1+X_2+\cdots+X_n$$

とおくと，Kolmogorov の不等式を $X_{n+1}, X_{n+2}, \cdots, X_{n+m}$ に適用して

$$P\left\{\max_{k=1}^{m}|S_{n+k}-S_n|>\varepsilon\right\} \leqq \frac{1}{\varepsilon^2}\sum_{k=1}^{m}V(X_{n+k}),$$

$$|S_{n+k}-S_{n+l}| \leqq |S_{n+k}-S_n|+|S_{n+l}-S_n|$$

であるから

$$\max_{1\leqq k,l\leqq m}|S_{n+k}-S_{n+l}| \leqq 2\max_{k=1}^{m}|S_{n+k}-S_n|.$$

したがって $v_n=V(X_n)$ とおいて

$$P\left\{\max_{1\leqq k,l\leqq m}|S_{n+k}-S_{n+l}|>2\varepsilon\right\} \leqq \frac{1}{\varepsilon^2}\sum_{k=1}^{\infty}v_{n+k}.$$

この ω 集合は m とともに増大して $\left\{\sup_{k,l}|S_{n+k}-S_{n+l}|>2\varepsilon\right\}$ に近づくから

$$P\left\{\sup_{k,l}|S_{n+k}-S_{n+l}|>2\varepsilon\right\} \leqq \frac{1}{\varepsilon^2}\sum_{k=1}^{\infty}v_{n+k}.$$

n が増大するとき，$\left\{\sup_{k,l}|S_{n+k}-S_{n+l}|>2\varepsilon\right\}$ は減少するから，

$$P\left\{\lim_{n\to\infty}\sup_{k,l}|S_{n+k}-S_{n+l}|>2\varepsilon\right\} \leqq \frac{1}{\varepsilon^2}\sum_{k=1}^{\infty}v_{n+k}, \qquad n=1,2,\cdots.$$

$\sum v_n<\infty$ により，右辺は $n\to\infty$ のとき0に収束し，

$$P\left\{\lim_{n\to\infty}\sup_{k,l}|S_{n+k}-S_{n+l}|>2\varepsilon\right\} = 0.$$

$\varepsilon\downarrow 0$ として

$$\lim_{n\to\infty}\sup_{k,l}|S_{n+k}-S_{n+l}| = 0 \text{ a.s.}$$

を得る．これは確率変数列 $\{S_n\}$ の概収束を意味する．∎

この定理は極めて有力であるが，$E(X_n)$，$V(X_n)$ が有限確定でないときには使えない．一般の場合に役立つ定理としてつぎの3級数定理がある．

定理 4.4 (Hinčin の3級数定理) $|X_n|\leqq 1$ か否かに応じて，$X_n'=X_n$，0 とする．すなわち

$$X_n' = X_n 1_{[-1,1]}(X_n).$$

$\sum X_n$ が概収束するための必要十分条件は

$$\sum_n E(X_n'),\quad \sum_n V(X_n'),\quad \sum_n P(X_n\neq X_n')$$

がすべて収束することである．

証明 十分性 $\{X_n\}$ の独立性から，$\{X_n'\}$ の独立性がでる (定理 3.14 (i))．前定理を $\{X_n'\}$ に適用して $\sum X_n'$ の概収束がでる．$\sum P(X_n\neq X_n')$ が収束するから，Borel-Cantelli の補題により，ほとんど確実に，ある番号以後 $X_n=X_n'$ となる．したがって $\sum X_n$ も概収束する．

必要性 $\sum X_n$ が概収束とする．もし $\sum P(X_n\neq X_n')$ が発散すれば，Borel の定理により，ほとんど確実に，無限に多くの n に対し

$$X_n \neq X_n' \quad \text{すなわち} \quad |X_n|>1$$

となり，$\sum X_n$ が概発散となる．したがって $\sum P(X_n\neq X_n')$ は収束する．これから Borel-Cantelli の補題によりほとんど確実にある番号以後 $X_n=X_n'$ となるから，$\sum X_n'$ が概収束することがわかる．この和を S とする．X_n'，$S_n\equiv\sum_{k=1}^{n}X_k'$，$S$ の特性関数をそれぞれ $\varphi_n(z)$，$\psi_n(z)$，$\psi(z)$ とし，X_n' の確率法則を μ_n とする．S_n は S に概収束し，$\{X_n'\}$ は独立であるから

$$|\psi(z)|^2 = \lim_{n\to\infty}|\psi_n(z)|^2 = \lim_{n\to\infty}\prod_{k=1}^{n}|\varphi_k(z)|^2 = \prod_{k=1}^{\infty}|\varphi_k(z)|^2.$$

$\psi(z)$ は S の特性関数であるから，$z=0$ のある近傍 U で0でない．したがって $z\in U$ で

$$\sum_n(1-|\varphi_n(z)|^2) < \infty.$$

$|\varphi_n(z)|^2$ は $\nu_n=\mu_n*\check{\mu}_n$ ($\check{\mu}_n$ は μ_n の裏返し) の特性関数であり，ν_n は0に関して対称な分布であり，しかも $[-2,2]$ に集中しているから，

$$\sum_n\int_{-2}^{2}(1-\cos z\xi)\,\nu_n(d\xi) < \infty.$$

十分小さい ξ に対し，$1-\cos\xi>\xi^2/3$ であることに注意すれば，十分小さい0の近傍 $V\subset U$ の中では

$$\sum_n\int_{-2}^{2}z^2\xi^2\nu_n(d\xi)<\infty \quad すなわち \quad \sum_n\int_{R^1}\xi^2\nu_n(d\xi)<\infty$$

となる．

$$M(\nu_n)=M(\mu_n)+M(\check{\mu}_n)=0,\qquad V(\nu_n)=V(\mu_n)+V(\check{\mu}_n)=2V(\mu_n)$$

により，上の式は $\sum_n V(\mu_n)<\infty$ を意味する．前定理により $\sum(X_n'-E(X_n'))$ は概収束である．$\sum X_n'$ が概収束であるから $\sum E(X_n')$ も収束しなければならない．∎

一般の確率変数を項とする級数については法則収束，確率収束，概収束は異なっている．しかし独立確率変数の場合にはこの3収束は同等であることを示そう．

定理 4.5 (P. Lévy の3収束同等定理)　$\{X_n\}$ を独立実確率変数列，$S_n=\sum_{k=1}^{n}X_k$ とするとき，つぎの3条件は同等である．

(i)　S_n の確率法則が収束する，

(ii)　S_n が確率収束する，

(iii)　S_n が概収束する．

証明　'(iii)$\Longrightarrow$(ii)$\Longrightarrow$(i)' は自明である．'(i)$\Longrightarrow$(ii)' を証明しよう．X_n の確率法則を μ_n とし，$\varphi_n=\mathscr{F}\mu_n$ とする．また

$$S_{mn}=\sum_{k=m+1}^{n}X_k\qquad(m<n)$$

の確率法則を μ_{mn}，$\varphi_{mn}=\mathscr{F}\mu_{mn}$ とする．仮定により μ_{1n} は $n\to\infty$ のときある分布 μ に収束するから，φ_{1n} は $\varphi\equiv\mathscr{F}\mu$ に収束する．しかも $\varphi_{1n}(z)$ の収束は有界な z 集合の上では一様である．$|\varphi(z)|$ は $z=0$ のある近傍 $|z|\leqq a$ で，ある正数 b より大きいから，十分大きい n に対し

$$|\varphi_{1n}(z)|\geqq\frac{b}{2}\qquad(|z|\leqq a).$$

$|z|\leqq a$ で $\varphi_{1n}(z)$ は $\varphi(z)$ に一様収束するから，$\varepsilon>0$ に対し $N(\varepsilon)$ を十分大きくとると

$$m>n>N(\varepsilon)\Longrightarrow|\varphi_{1n}(z)-\varphi_{1m}(z)|<\frac{b}{2}\varepsilon\qquad(|z|\leqq a).$$

$\varphi_{1m}=\varphi_{1n}\varphi_{nm}$ により

$$\begin{aligned}
&m>n>N(\varepsilon)\\
&\Longrightarrow\left|\int(1-e^{iz\xi})\mu_{nm}(d\xi)\right|=|1-\varphi_{nm}(z)|<\varepsilon\qquad(|z|\leqq a)\\
&\Longrightarrow\left|\frac{1}{2a}\int_{-a}^{a}dz\int(1-e^{iz\xi})\mu_{nm}(d\xi)\right|<\varepsilon\\
&\Longrightarrow\int\left(1-\frac{\sin a\xi}{a\xi}\right)\mu_{nm}(d\xi)<\varepsilon.
\end{aligned}$$

$\beta>0$ を適当に小さくとれば，この被積分関数は $\beta a^2\xi^2(1+a^2\xi^2)^{-1}$ より大きいから，上の不等式から

$$\int\frac{a^2\xi^2}{1+a^2\xi^2}\mu_{nm}(d\xi)<\frac{\varepsilon}{\beta},$$

したがって

$$\mu_{nm}([-\eta,\eta]^c)<\frac{\varepsilon}{\beta}\cdot\frac{1+a^2\eta^2}{a^2\eta^2}\qquad(\eta>0)$$

がでる．μ_{nm} は $S_{nm}=S_m-S_n$ の確率法則であるから，

$$m,n>N(\varepsilon)\Longrightarrow P\{|S_m-S_n|>\eta\}<\varepsilon(1+a^2\eta^2)\beta^{-1}a^{-2}\eta^{-2}$$

を意味する．これは $\{S_n\}$ が確率収束することを意味する．これで '(i)$\Longrightarrow$(ii)' が証明された．

残る所は '(ii)$\Longrightarrow$(iii)' の証明である．$\{S_n\}$ が確率収束するから，十分大きい n,m に対しては一様に

$$P\{|S_m-S_n|>\varepsilon\}<\frac{1}{2}.$$

Ottaviani の不等式により，$\varepsilon>0$ に対して n を十分大きくとれば，すべての m に対し

$$P\left\{\max_{k=1}^{m}|S_{n+k}-S_n|>2\varepsilon\right\}\leqq 2P\{|S_{n+m}-S_n|>\varepsilon\},$$

したがって

$$\begin{aligned}
&P\left\{\max_{1\leqq k,l\leqq m}|S_{n+k}-S_{n+l}|>4\varepsilon\right\}\\
&\leqq 4P\{|S_{n+m}-S_n|>\varepsilon\}\leqq 4\sup_m P\{|S_{n+m}-S_n|>\varepsilon\}.
\end{aligned}$$

これから定理 4.3 の証明の最後の部分の論法を用い，上の右辺が $n\to\infty$ のとき

0 に収束する (確率収束の仮定！) ことに注意すれば，$\{S_n\}$ の概収束がでる．▮

例題 4.2　(i)　$\{X_n\}$ が独立実確率変数列で，X_n の確率法則は区間 $[-1/n, 1/n]$ の上の一様分布とする．このとき $\sum X_n$ は概収束であるが，$\sum |X_n|$ は概発散であることを示せ．これは $\sum X_n$ の収束はほとんど確実に条件付収束であることを示している．

［ヒント］　定理 4.3 により $\sum X_n$ の概収束がでる．$\sum |X_n|$ の概発散の証明には $|X_n| \leqq 1$, $\sum E(|X_n|) = \infty$ に注意して 3 級数定理 (必要性の部分の対偶) を用いよ．

(ii)　$\{X_n\}$ が独立実確率変数列で，X_n の確率法則は Gauss 分布 $N(m_n, v_n)$ とする．$\sum X_n$ が概収束であるための必要十分条件は $\sum m_n$, $\sum v_n$ がともに収束することであることを示せ．

［ヒント］　十分性は定理 4.3 による．必要性をいうには

$$S_n = \sum_{k=1}^{n} X_k \longrightarrow S \text{ a.s.}$$

と仮定して $\sum m_n$, $\sum v_n$ の収束をいえばよい．仮定により S_n の確率法則は $N(M_n, V_n)$ $\left(M_n = \sum_1^n m_k,\ \ V_n = \sum_1^n v_k\right)$ である．

$$\exp\left\{-\frac{V_n}{2}z^2\right\} = |E(e^{izS_n})| \longrightarrow |E(e^{izS})| \qquad (n \to \infty).$$

$E(e^{izS})$ は十分小さい $z \neq 0$ で 0 とならないから，$\{V_n\}$ は収束しなければならない．これは $\sum v_n$ の収束を意味する．$Y_n = X_n - m_n$ とおくと，$EY_n = 0$, $\sum V(Y_n) = \sum v_n < \infty$ により，定理 4.3 を用いて $\sum Y_n$ は概収束．$\sum X_n$ は仮定により概収束であるから，$m_n = X_n - Y_n$ により $\sum m_n$ が収束する．

§4.3　中心値，散布度

$\{X_n\}$ を独立実確率変数列とし，X_n の確率法則がすべて Gauss 分布である (このことを X_n が **Gauss 分布に従う**ということもある) とし，

$$S_n = X_1 + X_2 + \cdots + X_n, \qquad n = 1, 2, \cdots$$

とおく．前節例題 (ii) によれば

$$S_n \text{ が概収束} \iff \{E(S_n)\},\ \{V(S_n)\} \text{ がともに収束}$$

となる．X_n が Gauss 分布に従うという条件をとり除いてこの事実に対応するこ

とを証明するのが本節の主要目的である．X_n が一般の場合には $E(S_n)$, $V(S_n)$ が定義できないから，まず平均値，分散に相当する概念で，任意の実確率変数に対して定義できる**中心値**，**散布度**を導入しよう．

X を任意の実確率変数とする．arctan は $(-\infty, \infty)$ から $(-\pi/2, \pi/2)$ への増大全単射であるから

$$E(\arctan(X-\gamma)) = 0$$

となる実数 γ が一意に定まる．これを $\gamma(X)$ であらわし，X の**中心値**という．明らかに

$$\gamma(X+a) = \gamma(X) + a \qquad (a \text{ は定数}),$$
$$\gamma(-X) = -\gamma(X).$$

また X の**散布度** $\delta(X)$ は

$$\delta(X) = -\log\left[\int\!\!\int e^{-|\xi-\eta|}\mu(d\xi)\,\mu(d\eta)\right] \in [0, \infty) \qquad (\mu = P^X)$$

で定義される．明らかに

$$\delta(X+a) = \delta(X), \qquad \delta(-X) = \delta(X).$$

$X\ (\in L^1)$ の平均値 $E(X)$ は $E(X-m) = 0$ をみたす m であるから，中心値は平均値に対応し，一般の X に対して定義される．

$X\ (\in L^2)$ の分散 $V(X)$ は

$$V(X) = \inf_{-\infty < m < \infty} E((X-m)^2) \qquad (\because\ E((X-m)^2) = V(X) + (m - EX)^2)$$

で与えられるから，X の値のばらつきを示している．$\delta(X)$ も $X\ (\in L^0)$ の値のばらつきを表わしていることはつぎの事実からわかる．

$$\mu[a-l, a+l]^2 = \int\!\!\int_{|\xi-a|, |\eta-a| \leqq l} \mu(d\xi)\,\mu(d\eta) \leqq \int\!\!\int_{|\xi-\eta| \leqq 2l} \mu(d\xi)\,\mu(d\eta),$$
$$e^{-\delta(X)} = \int\!\!\int e^{-|\xi-\eta|}\mu(d\xi)\,\mu(d\eta) \geqq e^{-2l}\int\!\!\int_{|\xi-\eta| \leqq 2l} \mu(d\xi)\,\mu(d\eta),$$
$$\mu[a-l, a+l] \leqq e^{l-\delta(X)/2}.$$

したがって $\delta(X)$ が極めて大きいときには長さ一定 $(2l)$ の区間の中に X が入る確率は一様に極めて小さい．また

$$e^{-\delta(X)} = \int\!\!\int e^{-|\xi-\eta|}\mu(d\xi)\,\mu(d\eta)$$

$$\leqq \int (e^{-l}\mu([-l+\eta, l+\eta]^c)+\mu[-l+\eta, l+\eta])\,\mu(d\eta)$$
$$\leqq e^{-l}+\sup_a \mu[a-l, a+l].$$

したがって十分大きい l $(e^{-l}<\varepsilon)$ をきめたとき，長さ $2l$ の区間の中に X が入る確率が一様に小さい $(<\varepsilon)$ ならば，$\delta(X)$ は極めて大きい $(>\log(1/2\varepsilon))$．また

$$\delta(X)=0 \iff X=a \text{ a.s.} \quad (a\ はある定数)$$

がなりたつ．実際 X が定数のときには $\delta(X)=0$ は明らか．$\delta(X)=0$ ならば，$\mu=P^X$ に対し

$$\iint e^{-|\xi-\eta|}\mu(d\xi)\,\mu(d\eta)=1.$$

被積分関数 $\leqq 1$ であるから，ある a に対し

$$\int e^{-|\xi-a|}\mu(d\xi)=1.$$

したがって

$$\mu\{a\}=1 \quad すなわち \quad X=a \text{ a.s.}$$

さらに

$$\delta(X_n)\longrightarrow 0 \iff 適当な定数列\ \{a_n\}\ に対し\ X_n-a_n\longrightarrow 0 \text{ i.p.}$$

が証明される．実際 $X_n-a_n\to 0$ i.p. ならば，$\delta(X_n-a_n)\to 0$ は容易にわかる．これから $\delta(X_n)\to 0$ がでる．$\delta(X_n)\to 0$ とせよ．

$$\iint e^{-|\xi-\eta|}\mu_n(d\xi)\,\mu_n(d\eta)\longrightarrow 1 \qquad (\mu_n=P^{X_n}),$$
$$\sup_\eta \int e^{-|\xi-\eta|}\mu_n(d\xi)\longrightarrow 1.$$

したがって $\{a_n\}$ が存在して

$$\int e^{-|\xi-a_n|}\mu_n(d\xi)\longrightarrow 1 \quad すなわち \quad E(e^{-|X_n-a_n|})\longrightarrow 1.$$

Bienaymé の不等式により，$\varepsilon>0$ に対し

$$P(|X_n-a_n|>\varepsilon)\leqq \frac{1}{1-e^{-\varepsilon}}E(1-e^{-|X_n-a_n|})\longrightarrow 0.$$

これは '$X_n-a_n\to 0$ i.p.' を示す．

以上の事実により，$\delta(X)$ が X のばらつきの程度をあらわしていることがわかる．

さて $\{X_n\}\subset L^2$ が独立のときには，$S_n=\sum_{k=1}^n X_k$ に対し，

$$V(S_n)=\sum_{k=1}^n V(X_k)$$

は n とともに増大する．散布度 δ も同じ性質をもつ．

定理 4.6　$\{X_n\}$ が任意の独立実確率変数のとき，$S_n=\sum_1^n X_k$ の散布度 $\delta(S_n)$ は n とともに増大する．

証明　まず

$$\iint e^{-|\xi-\eta|}\mu(d\xi)\,\mu(d\eta)$$
$$=\int e^{-|\xi|}(\mu*\check{\mu})(d\xi) \qquad (\check{\mu}\ は\ \mu\ の裏返し)$$
$$=\frac{1}{\pi}\iint \frac{e^{i\xi x}}{1+x^2}dx\,(\mu*\check{\mu})(d\xi)$$
$$=\frac{1}{\pi}\int\frac{dx}{1+x^2}\int e^{i\xi x}(\mu*\check{\mu})(d\xi)$$
$$=\frac{1}{\pi}\int\frac{dx}{1+x^2}|\mathscr{F}\mu(x)|^2$$

を注意しておく．X_n の特性関数を φ_n とすれば，S_n の特性関数は $\prod_1^n \varphi_k$ である．したがって上の注意により

$$e^{-\delta(S_n)}=\frac{1}{\pi}\int\frac{dx}{1+x^2}\left|\prod_{k=1}^n\varphi_k(x)\right|^2.$$

$|\varphi_k(x)|\leqq 1$ であるから，n の増大に応じて右辺は減少し，したがって $\delta(S_n)$ は増大する．▌

$\{X_n\}\subset L^2$ のときには $V(S_n)=V(S_{n+1})$ ならば $V(X_{n+1})=0$ したがって 'X_{n+1} =定数 a.s.' であった．このことは一般の場合に δ に対してもいえる．$|\varphi_{n+1}(x)|\leqq 1$ により $\delta(S_n)=\delta(S_{n+1})$ ならば

$$\left|\prod_1^n\varphi_k(x)\right|^2=\left|\prod_1^{n+1}\varphi_k(x)\right|^2 \quad \text{a.e.}\ (\text{Lebesgue}\ 測度について).$$

φ_k $(k=1,2,\cdots,n)$ は連続で左辺は $x=0$ のある近傍 $|x|\leqq a$ で 1 に近いから

$$|\varphi_{n+1}(x)|^2=1 \qquad (|x|\leqq a).$$

X_{n+1} の確率法則を μ とすれば，$|\varphi_{n+1}|^2$ は $\mathcal{F}\tilde{\mu}$ $(\tilde{\mu}=\mu*\check{\mu})$ に等しい．$\tilde{\mu}$ は対称であるから，上の式から $|x|\leqq a$ で

$$\int \cos x\xi\,\tilde{\mu}(d\xi) = 1 \quad すなわち \quad \int (1-\cos x\xi)\,\tilde{\mu}(d\xi) = 0.$$

x について $[-a, a]$ で平均して

$$\int \left(1-\frac{\sin a\xi}{a\xi}\right)\tilde{\mu}(d\xi) = 0.$$

被積分関数は $\xi\neq 0$ では正であるから，$\tilde{\mu}$ は 0 に集中している．これから μ はある1点に集中していることがわかる．これは X_{n+1} がほとんど確実にある定数に等しいことを示している．

さてもとの問題にもどりつぎの定理を証明する．

定理 4.7 $\{X_n\}$ を一般の独立実確率変数列とし，

$$S_n = X_1+X_2+\cdots+X_n, \qquad n = 1, 2, \cdots$$

とする．

(i) 適当な a_n があって $\{S_n-a_n\}$ が概収束 $\Longleftrightarrow$ $\delta(S_n)$ が収束，

(ii) $\{S_n\}$ が概収束 $\Longleftrightarrow$ $\gamma(S_n)$, $\delta(S_n)$ がともに収束．

証明 $\delta(S_n)$ は増大するから

$$\delta_\infty = \lim_{n\to\infty}\delta(S_n)$$

が確定する．

(i) $T_n=S_n-a_n\to T$ a.s. とする．前定理の証明で注意したように

$$\delta(S_n) = \delta(T_n) = -\log\left(\frac{1}{\pi}\int\frac{dx}{1+x^2}|E(e^{ixT_n})|^2\right).$$

$n\to\infty$ として

$$\delta_\infty = -\log\left(\frac{1}{\pi}\int\frac{dx}{1+x^2}|E(e^{ixT})|^2\right) = \delta(T) < \infty.$$

これで'$\Longrightarrow$'が示された．

逆に $\delta_\infty<\infty$ とせよ．X_n の確率法則，特性関数をそれぞれ μ_n, φ_n とする．

$$\delta_\infty = \lim_{n\to\infty}\delta(S_n) = -\log\left(\frac{1}{\pi}\int\frac{dx}{1+x^2}\prod_{n=1}^{\infty}|\varphi_n(x)|^2\right).$$

$\delta_\infty<\infty$ であるから

$$\int\frac{dx}{1+x^2}\prod_{n=1}^{\infty}|\varphi_n(x)|^2 > 0.$$

したがって (Lebesgue) 測度正のある集合 A の上で

$$\prod_{n=1}^{\infty}|\varphi_n(x)|^2 > 0,$$

すなわち

$$\sum_{n=1}^{\infty}(1-|\varphi_n(x)|^2) < \infty$$

となる．正数 C と測度正の $B\subset A$ を適当にとって

$$\sum_n (1-|\varphi_n(x)|^2) < C \qquad (x\in B)$$

とすることができる．B の測度 $\leqq 1$ と仮定してよい．$|\varphi_n(x)|^2$ は $\tilde{\mu}_n=\mu_n*\check{\mu}_n$ の特性関数であるから，上の式から

$$\sum_n\int(1-\cos x\xi)\,\tilde{\mu}_n(d\xi) < C \qquad (x\in B),$$

$$\sum_n\int\int_B(1-\cos x\xi)\,dx\,\tilde{\mu}_n(d\xi) < C.$$

しかるに

$$(*) \qquad \int_B(1-\cos x\xi)\,dx \geqq C_1\frac{\xi^2}{1+\xi^2} \qquad (C_1\text{ は適当な正定数})$$

である (証明は後述) から，

$$\sum_n\int\frac{\xi^2}{1+\xi^2}\tilde{\mu}_n(d\xi) < C_2 \equiv \frac{C}{C_1}.$$

$\tilde{\mu}_n=\mu_n*\check{\mu}_n$ であるから，上の式は

$$\sum_n\int\int\frac{(\xi-\eta)^2}{1+(\xi-\eta)^2}\mu_n(d\xi)\,\mu_n(d\eta) < C_2$$

となる．これから

$$\sum_n\inf_\eta\int\frac{(\xi-\eta)^2}{1+(\xi-\eta)^2}\mu_n(d\xi) < C_2.$$

ゆえに c_n を適当にとって

$$\sum_n\int\frac{(\xi-c_n)^2}{1+(\xi-c_n)^2}\mu_n(d\xi) < C_2,$$

$$\sum_n \int_{|\xi-c_n|\leq 1}(\xi-c_n)^2\mu_n(d\xi)<\infty, \quad \sum_n \int_{|\xi-c_n|>1}\mu_n(d\xi)<\infty.$$

$|X_n-c_n|\leq 1$ か否かに応じて $Y_n=X_n-c_n$, 0 とおけば，上の不等式は

$$\sum_n V(Y_n)\leq \sum_n E(Y_n^2)<\infty, \quad \sum_n P(Y_n\neq X_n-c_n)<\infty.$$

第1の不等式から定理4.3により $\sum_n (Y_n-E(Y_n))$ が概収束，第2の不等式から，ほとんど確実にある番号以後 $Y_n=X_n-c_n$ となることがわかる．したがって $\sum (X_n-c_n-E(Y_n))$ は概収束である．これは

$$S_n-\sum_{k=1}^n (c_k+E(Y_k)), \quad n=1,2,\cdots$$

が概収束することを意味する．

残しておいた不等式 (*) を証明すれば (i) の証明が完了する．

$$f(\xi)=\frac{1+\xi^2}{\xi^2}\int_B(1-\cos x\xi)\,dx=\frac{1+\xi^2}{\xi^2}\left(|B|-\int\cos x\xi\, 1_B(x)\,dx\right).$$

ここで $|B|$ は B の Lebesgue 測度をあらわす．$|B|>0$ であるから $f(\xi)>0$．$|B|<\infty$ であるから右辺の積分は $|\xi|\to\infty$ のとき 0 に収束し (Riemann-Lebesgue の定理 !)，したがって $f(\xi)\to|B|>0$ $(|\xi|\to\infty)$．また $\xi\to 0$ のときには

$$f(\xi)=\int_B(1-\cos x\xi)\frac{1+\xi^2}{\xi^2}dx\longrightarrow\int_B\frac{1}{2}x^2dx>0.$$

これから $C_1=\inf_\xi f(\xi)$ が正となり，この C_1 に対し (*) がなりたつ．

(ii) $\{S_n\}$ が S に概収束すれば，(i) により $\delta(S_n)$ が収束．$\gamma(S_n)\to\gamma(S)$ を証明しよう．$\{\gamma(S_n)\}$ の任意の極限点を $l\in[-\infty,\infty]$ とするとき，$l=\gamma(S)$ をいえばよい．$\gamma(S_{p(n)})\to l$ とする．$\{p(n)\}$ は $\{n\}$ の部分列である．

$$\arctan(S_{p(n)}-\gamma(S_{p(n)}))\longrightarrow\arctan(S-l)$$
$$(\arctan(\pm\infty)=\pm\pi/2 \text{ とする})$$

に注意し，有界収束定理により

$$E(\arctan(S-l))=0,$$

したがって $l=\gamma(S)$ である．

逆に $\{\gamma(S_n)\}$, $\{\delta(S_n)\}$ が収束するとせよ．$\{\delta(S_n)\}$ の収束から，(i) により適当な a_n があって，$\{S_n-a_n\}$ が概収束する．したがって，上に証明したことから $\{\gamma(S_n-a_n)\}$ が収束する．仮定により $\{\gamma(S_n)\}$ が収束するから，$\{a_n\equiv\gamma(S_n)-\gamma(S_n-a_n)\}$ が収束する．$\{S_n-a_n\}$ が概収束するから，$\{S_n\}$ も概収束する．∎

上の (ii) の証明で，一般に

$$S_n\longrightarrow S \text{ a.s.} \Longrightarrow \gamma(S_n)\longrightarrow\gamma(S)$$

がなりたつことを示した．このことを利用して (i) の $\{a_n\}$ として $\{\gamma(S_n)\}$ をとれることがわかる．実際 $\{S_n-a_n\}$ が概収束すれば，$\gamma(S_n-a_n)=\gamma(S_n)-a_n$ により，$\{\gamma(S_n)-a_n\}$ が収束する．したがって $\{S_n-\gamma(S_n)\}$ も概収束する．

例題 4.3 (i) $\{X_k\}\subset L^2$ が独立とし，$\sum_n V(X_n)<\infty$ とする．このとき

$$\sum_n X_n \text{ が概収束} \Longleftrightarrow \sum_n E(X_n) \text{ が収束}.$$

［ヒント］ $\sum V(X_n)<\infty$ により定理4.3を用いて $\sum_n(X_n-EX_n)$ の概収束がでる．

(ii) '$X_n\to X$ i.p.' ならば $\gamma(X_n)\to\gamma(X)$，$\delta(X_n)\to\delta(X)$ となることを示せ．

［ヒント］ $\{\gamma(X_n)\}$ の任意の極限点 $l\in[-\infty,\infty]$ が $\gamma(X)$ に等しいことをいえば $\gamma(X_n)\to\gamma(X)$ がでる．$\gamma(X_{p(n)})\to l$ とする．仮定により '$X_{p(n)}\to X$ i.p.' であるから，$\{X_{p(n)}\}$ の適当な部分列 $\{X_{q(n)}\}$ が X に概収束する．したがって $\gamma(X_{q(n)})\to\gamma(X)$．これから $l=\gamma(X)$ を得る．'$X_n\to X$ i.p.' により，$n\to\infty$ のとき

$$e^{-\delta(X_n)}=\frac{1}{\pi}\int\frac{dz}{1+z^2}|E(e^{izX_n})|^2$$
$$\longrightarrow\frac{1}{\pi}\int\frac{dz}{1+z^2}|E(e^{izX})|^2=e^{-\delta(X)}.$$

(iii) 1次元の分布 μ に対して $\gamma(\mu)$ を

$$\int\arctan(\xi-\gamma)\,\mu(d\xi)=0$$

をみたす γ と定義し，$\delta(\mu)$ を

$$\delta(\mu)=-\log\left(\int\int e^{-|\xi-\eta|}\mu(d\xi)\,\mu(d\eta)\right)$$

で定義する．$\mu_n\to\mu$ のとき $\gamma(\mu_n)\to\gamma(\mu)$，$\delta(\mu_n)\to\delta(\mu)$ を証明せよ．

［ヒント］ $\{\gamma(\mu_n)\}$ の任意の極限点 l が $\gamma(\mu)$ に等しいことを示して，$\gamma(\mu_n)\to\gamma(\mu)$ をだす．この際 $|\arctan x-\arctan y|\leq|x-y|$ に注意せよ．$\delta(\mu_n)\to\delta(\mu)$ をだすには，上の2重積分が

$$\frac{1}{\pi}\int\frac{dx}{1+x^2}|(\mathscr{F}\mu)(x)|^2$$

に等しいことに注意する.

(iv) 上の (iii) を用いて, 定理 2.11 の条件がなりたつならば, $\delta(\mu), \gamma(\mu)$ ($\mu \in \mathscr{M}$) はともに有界であることを示せ.

[ヒント] $\delta(\mu_n)\to\infty$ となる $\mu_n \in \mathscr{M}$ があれば, その適当な部分列の極限分布 μ に対しては $\delta(\mu)=\infty$ となり矛盾. $\gamma(\mu)$ についても同様.

§4.4 独立確率変数の級数の概発散

$\{X_n\}$ を独立確率変数列とし

$$S_n = \sum_{k=1}^{n} X_k, \qquad n = 1, 2, \cdots$$

とおくと, 前節でのべたようにつぎの3通りの場合が可能である.

(C) $\{\delta(S_n)\}$, $\{\gamma(S_n)\}$ がともに収束のときには, $\{S_n\}$ が概収束,

(D_c) $\{\delta(S_n)\}$ が収束, $\{\gamma(S_n)\}$ が発散のときには, $\{S_n\}$ は概発散であるが, $\{S_n-\gamma(S_n)\}$ が概収束,

(D) $\{\delta(S_n)\}$ が発散のときには, すべての実数列 $\{a_n\}$ に対し $\{S_n-a_n\}$ が概発散.

(C) のときには $\sum X_n$ は**収束型**, (D_c) のときには収束型に帰着される発散型または簡単に**擬発散型**, (D) のときには**真発散型**とよぶことにしよう. ここで発散という言葉を用いたが, それは収束しないという意味であって, 絶対値が限りなく大きくなって行くとか, 有界な範囲の中で振動していて収束しないとか, 何度でも有界な範囲にもどって来てそこからでて遠くへ行くというようないろいろの型が考えられ, これらの型がそれぞれある確率でおこるということもあり得るかもしれない.

擬発散型 (D_c) の場合には $\{S_n-\gamma(S_n)\}$ が概収束であるから,

$$S_n = (S_n-\gamma(S_n))+\gamma(S_n), \qquad n = 1, 2, \cdots$$

の発散の状況は $\{\gamma(S_n)\}$ の発散の状況で定まるので, 問題はない.

真発散型 (D) の場合を考えて見よう. このときには $\delta(S_n)\to\infty$ $(n\to\infty)$ であるから, $\{S_n\}$ のある部分列 $\{S_{p(n)}\}$ を適当にとって

$$\sum_{n=1}^{\infty} e^{-\delta(S_{p(n)})/2} < \infty \qquad (\text{例えば } \delta(S_{p(n)})>n)$$

とすることができる. $\{a_n\}$ をどのように定めても

$$P\{|S_n-a_n|\leqq l\} \leqq e^{l-\delta(S_n)/2} \qquad (\text{前節参照})$$

であるから,

$$\sum_n P\{|S_{p(n)}-a_{p(n)}|\leqq l\} = e^l \sum_n e^{-\delta(S_{p(n)})/2} < \infty.$$

Borel-Cantelli の補題により,

$$P\{|S_{p(n)}-a_{p(n)}|>l \text{ f.e.}\} = 1,$$

したがって

$$P\left\{\limsup_{n\to\infty}|S_n-a_n|\geqq l\right\} = 1,$$

$l\to\infty$ として

$$P\left\{\limsup_{n\to\infty}|S_n-a_n|=\infty\right\} = 1.$$

かくしてつぎの定理を得る.

定理 4.8 真発散型のときには, すべての実数列 $\{a_n\}$ に対し

$$\limsup_{n\to\infty}|S_n-a_n| = \infty \text{ a.s.}$$

この定理により $\{S_n-a_n\}$ がある定まった有限区間 I の中に封じこめられるという確率は0であることがわかる. しかし $\{S_n-a_n\}$ が I に何度でも入ってくるという確率は必ずしも0ではない. 例えば

$$P\{X_n=-1\} = P\{X_n=1\} = \frac{1}{2}, \qquad n = 1, 2, \cdots$$

のときには,

$$E(e^{izX_n}) = \cos z, \qquad E(e^{izS_n}) = (\cos z)^n$$

であるから, 有界収束定理により

$$e^{-\delta(S_n)} = \frac{1}{\pi}\int\frac{(\cos x)^{2n}}{1+x^2}dx \longrightarrow 0 \quad \text{したがって} \quad \delta(S_n) \longrightarrow \infty$$

となり, $\sum X_n$ は真発散型である. したがって上の定理により

$$\limsup_{n\to\infty}|S_n| = \infty \text{ a.s.}$$

しかし'ほとんど確実に $S_n=0$ i.o.' となることが §4.7 で証明される.

もし $\delta(S_n)$ が $n\to\infty$ のとき急激に増大して

$$\sum_n e^{-\delta(S_n)/2} < \infty$$

であるならば，上と同様の論法で

$$\lim_{n\to\infty} |S_n - a_n| = \infty \text{ a.s.}$$

が証明される．

例題 4.4 (i) $\{X_n\}$ が独立確率変数列で，X_n が Gauss 分布 $N(m_n, v_n)$ に従うとき，$\sum X_n$ が収束型，擬発散型，真発散型であるための必要十分条件を $\sum m_n$, $\sum v_n$ の収束，発散の言葉であらわせ．

［ヒント］ $S_n = \sum_{k=1}^n X_k$ とすると $\gamma(S_n) = \sum_1^n m_k$,

$$e^{-\delta(S_n)} = \frac{1}{\pi}\int \frac{dx}{1+x^2} \exp\left\{-x^2 \sum_1^n v_k\right\},$$

$$\lim_n \delta(S_n) < \infty \iff \sum_1^\infty v_k < \infty.$$

(ii) $\{X_n\}$ が独立確率変数列で，X_n の確率法則はすべて同じ μ であるとする．μ が1点に集中していないならば，$\sum X_n$ は真発散型であることを示せ．

［ヒント］ $S_n = \sum_1^n X_k$ とすると

$$e^{-\delta(S_n)} = \frac{1}{\pi}\int \frac{dx}{1+x^2} |\varphi(x)|^{2n} \qquad (\varphi = \mathscr{F}\mu).$$

$|\varphi(x)|$ が可算個の点を除いて1に等しくないことをいえば，$n\to\infty$ のとき右辺の積分は0に収束し，$\delta(S_n)\to\infty$ となる．もし

$$|\varphi(x_1)| = |\varphi(x_2)| = 1, \qquad x_1, x_2 \neq 0, \qquad x_1 \neq x_2$$

ならば

$$\int \cos x_1\xi \tilde{\mu}(d\xi) = \int \cos x_2\xi \tilde{\mu}(d\xi) = 1 \qquad (\tilde{\mu} = \mu * \check{\mu}).$$

ゆえに

$$\tilde{\mu} \text{ の台} \subset \{2\pi n/x_i, n=0, \pm 1, \pm 2, \cdots\} \qquad (i=1,2).$$

μ が1点に集中していないならば，$\tilde{\mu}$ の台は少なくとも1点 $\xi_0 \neq 0$ を含む．ゆえに $\xi_0 = 2\pi n_i/x_i$ $(i=1,2)$．当然 $n_1, n_2 \neq 0$．したがって x_1/x_2 は有理数．これから $|\varphi(x)|=1$ となる x は可算個しかあり得ないことがわかる．

§4.5 大数の強法則

$\{X_n\}$ を独立実確率変数列とし，

$$S_n = \sum_{k=1}^n X_k, \qquad n = 1, 2, \cdots$$

とする．$\sum X_n$ が真発散型のときには，任意の定数列 $\{a_n\}$ に対して

$$\limsup_{n\to\infty} |S_n - a_n| = \infty \text{ a.s.}$$

であることは前節で証明した．しかし ∞ に発散する正の定数列 $\{b_n\}$ に対しては，$\{a_n\}$ を適当に選んで

$$\limsup_{n\to\infty} \frac{|S_n - a_n|}{b_n} = 0 \quad \text{すなわち} \quad \lim_{n\to\infty} \frac{S_n - a_n}{b_n} = 0 \text{ a.s.}$$

となる可能性がある．このときには $\{|S_n - a_n|\}$ は $\{b_n\}$ と比較すればはるかに小さい．本節では $\{S_n\}$ に関するこのような型の評価を論じよう．

まず準備として次の補題を証明する．

補題 4.1 (Kronecker の補題) $\{x_n\}$ が実数列，$\{b_n\}$ が ∞ に発散する増加正数列とするとき，

$$\sum_n \frac{x_n}{b_n} \text{ が収束} \Longrightarrow \frac{1}{b_n}\sum_{k=1}^n x_k \longrightarrow 0.$$

証明 $n = 0, 1, 2, \cdots, \infty$ に対し

$$s_0 = 0, \qquad s_n = \sum_{k=1}^n \frac{x_k}{b_k}$$

とおくと

$$x_n = b_n(s_n - s_{n-1}).$$

Abel の部分和公式を用いて

$$\frac{1}{b_n}\sum_{k=1}^n x_k = \frac{1}{b_n}\sum_{k=1}^n b_k(s_k - s_{k-1}) = s_n - \frac{1}{b_n}\sum_{k=1}^n s_{k-1}(b_k - b_{k-1}) \qquad (b_0 = 0),$$

$$b_k - b_{k-1} \geqq 0, \qquad \sum_{k=1}^n (b_k - b_{k-1}) = b_n \longrightarrow \infty, \qquad s_n \longrightarrow s_\infty$$

により，

$$\lim_{n\to\infty} \frac{1}{b_n}\sum_{k=1}^n x_k = s_\infty - s_\infty = 0. \quad \blacksquare$$

定理 4.9 $\{X_n\} \subset L^2$ が独立で，$\{b_n\}$ が ∞ に発散する増加正数列とする．

$$\sum_n \frac{V(X_n)}{b_n^{\,2}} < \infty \Longrightarrow \frac{S_n - E(S_n)}{b_n} \longrightarrow 0 \quad (n\to\infty) \quad \text{a.s.}$$

証明 X_n のかわりに $X_n - E(X_n)$ を考えることにより，$E(X_n)=0$（したがって $E(S_n)=0$）と仮定して一般性を失わない．独立確率変数列

$$Y_n = \frac{X_n}{b_n}, \qquad n = 1, 2, \cdots$$

は

$$EY_n = 0, \qquad \sum_n V(Y_n) = \sum_n \frac{V(X_n)}{b_n^{\,2}} < \infty$$

をみたすから，定理4.3(Kolmogorov の定理)により $\sum X_n/b_n$ は概収束する．したがって上の Kronecker の補題により $\{S_n/b_n\}$ はほとんど確実に0に収束する．∎

独立確率変数列 $\{X_n\}\subset L^2$ の分散の和 $\sum V(X_n)$ が収束するときには，$S_n - E(S_n)$ は収束するから，任意の正数列 $b_n\to\infty$ に対し $(S_n - E(S_n))/b_n \to 0$．$\sum V(X_n)$ が発散するときには，上の定理を応用してつぎの定理を得る．

定理 4.10 $\{X_n\}\subset L^2$ が独立で，$\sum V(X_n)$ が発散すれば任意の $\varepsilon>0$ に対して

$$\frac{S_n - E(S_n)}{V(S_n)^{1/2+\varepsilon}} \longrightarrow 0 \quad \text{a.s.}$$

証明 $V(S_n)\to\infty$ であるから，ある $V(S_m)>0$．したがって必要ならば $X_1' = S_m$, $X_2' = X_{m+1}$, $X_3' = X_{m+2}$, $\cdots$ を考えることにより $V(S_1)>0$ と仮定してよい．$V(S_n) = \sum_{k=1}^{n} V(X_k) \to \infty$ $(n\to\infty)$ であるから，前定理で $b_n = V(S_n)^{1/2+\varepsilon}$ とおいて $\sum V(X_n)/b_n^{\,2}$ が収束することをいえばよい．$V(X_n) = V(S_n) - V(S_{n-1})$ に注意して

$$\sum_{n=2}^{\infty} \frac{V(X_n)}{b_n^{\,2}} \leqq \sum_{n=2}^{\infty} \int_{V(S_{n-1})}^{V(S_n)} \frac{dx}{x^{1+2\varepsilon}} = \int_{V(S_1)}^{\infty} \frac{dx}{x^{1+2\varepsilon}} < \infty. \qquad ∎$$

上の定理の特別の場合としてつぎの定理が得られる．

定理 4.11 (Kolmogorov の大数の強法則) $\{X_n\}\subset L^2$ が独立で，$\{V(X_n)\}$ が有界ならば

$$\frac{S_n - E(S_n)}{n} \longrightarrow 0 \quad \text{a.s.}$$

証明 $\sum V(X_n)$ が収束ならば定理4.3により $S_n - E(S_n)$ は概収束するから，

定理の結論は自明である．$\sum V(X_n)$ が発散するときには前定理で $\varepsilon=1/2$ とおいて

$$\frac{S_n - E(S_n)}{V(S_n)} \longrightarrow 0 \quad \text{a.s.}$$

$V(X_n)$ は有界であるから，$a = \sup V(X_n) < \infty$ とおいて

$$V(S_n) = \sum_{k=1}^{n} V(X_n) \leqq na$$

となるから，'$(S_n - E(S_n))/n \to 0$ a.s.' が得られる．∎

つぎの定理では X_n がすべて同じ確率法則をもつという点では上の定理よりも仮定は強いが，そのかわり $\{X_n\}\subset L^2$ の仮定をゆるめて $\{X_n\}\subset L^1$ としている点に注意すべきである．

定理 4.12 (同分布の場合の大数の強法則) $\{X_n\}\subset L^1$ が独立で同じ分布 μ に従うとする．このとき

$$\frac{S_n}{n} \longrightarrow M(\mu) \quad \text{a.s.}$$

証明 $|X_n|\leqq n$ か否かに応じて $X_n' = X_n, 0$ とおき，

$$S_n' = \sum_{k=1}^{n} X_k'$$

とする．$|X_n|$ の確率法則は n に無関係であるから，これを ν とおく．

$$\begin{aligned}
\sum_{n=1}^{\infty} P\{X_n \neq X_n'\} &= \sum_{n=1}^{\infty} P\{|X_n|>n\} = \sum_{n=1}^{\infty} \nu(n,\infty) \\
&\leqq \int_0^\infty \nu(y,\infty)\,dy = \int_0^\infty \left(\int_y^\infty \nu(dx)\right) dy \\
&= \int_0^\infty \left(\int_0^x dy\right) \nu(dx) = \int_0^\infty x\nu(dx) \\
&= E(|X_1|) < \infty.
\end{aligned}$$

ゆえに Borel-Cantelli の補題により，

$$P\{X_n = X_n' \ \text{f.e.}\} = 1,$$

したがって

$$\frac{S_n'}{n} \longrightarrow M(\mu) \quad \text{a.s.}$$

がいえれば，定理の証明が完了する．

$$\frac{1}{n}E(S_n') = \frac{1}{n}\sum_{k=1}^{n} E(X_k'), \qquad E(X_k') = \int_{[-k,k]} x\mu(dx) \longrightarrow M(\mu)$$

であるから，$E(S_n')/n \to M(\mu)$．したがって

$$\frac{1}{n}(S_n' - E(S_n')) \longrightarrow 0 \text{ a.s.}$$

を証明してもよい．このためには定理 4.9 により（$|X_n'| \leqq n$ により $X_n' \in L^2$ となることに注意），

$$\sum_n \frac{V(X_n')}{n^2} < \infty$$

をいえば十分である．

$$\begin{aligned}
\text{左辺} &\leqq \sum_n \frac{E((X_n')^2)}{n^2} \leqq 4\sum_n \frac{E((X_n')^2)}{(n+1)^2} \\
&\leqq 4\sum_n \frac{1}{(n+1)^2}\int_{[0,n]} x^2\nu(dx) \qquad (\nu \text{ は上で導入した}) \\
&\leqq 4\int_0^\infty \frac{dy}{y^2}\int_{[0,y]} x^2\nu(dx) \\
&= 4\int_0^\infty \left(\int_x^\infty \frac{dy}{y^2}\right) x^2\nu(dx) = 4\int_0^\infty x\nu(dx) \\
&= 4\int_{-\infty}^\infty |x|\mu(dx) < \infty.
\end{aligned}$$

∎

例題 4.5 (i) $\varphi(t)$ を (a,∞) で定義された増大正値関数で（a はある正定数）

$$\varphi(t) \longrightarrow \infty \qquad (t \to \infty)$$

とし

$$\int_a^\infty \frac{dt}{\varphi(t)^2} < \infty \qquad (a > 0)$$

とする．（上の積分はある $a>0$ で有限ならば，任意の a に対して有限である．）このとき $V(S_n)^{1/2+\epsilon}$ のかわりに $\varphi(V(S_n))$ として定理 4.10 がなりたつことを示せ．

［ヒント］ 定理 4.10 の証明にならって

$$\sum_{n=2}^\infty \frac{V(X_n)}{\varphi(V(S_n))^2} \leqq \sum_{n=2}^\infty \int_{V(S_{n-1})}^{V(S_n)} \frac{dt}{\varphi(t)^2} = \int_{V(S_1)}^\infty \frac{dt}{\varphi(t)^2}.$$

(ii) 上の問題 (i) において

$$\varphi(t) = \sqrt{t(\log t)^{1+\epsilon}}$$

とおいて，定理 4.10 の仮定のもとで，より精密な結論

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)(\log V(S_n))^{1+\epsilon}}} \longrightarrow 0 \text{ a.s.}$$

がでることを示せ．また

$$\varphi(t) = \sqrt{t \log t \log^2 t \log^3 t \cdots \log^{k-1} t (\log^k t)^{1+\epsilon}}$$

とおくと，上の評価はいくらでも精密化される．ここで $\log^k t$ は

$$\log^1 t = \log t, \qquad \log^{k+1} t = \log(\log^k t)$$

によって定義される．

(iii) 前問を利用して定理 4.11 の仮定のもとで

$$\frac{S_n - E(S_n)}{\sqrt{n \log n \log^2 n \cdots \log^{k-1} n (\log^k n)^{1+\epsilon}}} \longrightarrow 0 \text{ a.s.}$$

がなりたつことを示せ．

(iv) $$\lim_{n\to\infty} \int_0^1\int_0^1 \cdots \int_0^1 \frac{x_1{}^q + x_2{}^q + \cdots + x_n{}^q}{x_1{}^p + x_2{}^p + \cdots + x_n{}^p} dx_1 dx_2 \cdots dx_n = \frac{p+1}{q+1}$$

$$(0 < p < q < \infty)$$

を証明せよ．

［ヒント］ $X_1, X_2, \cdots$ が独立で，すべて $[0,1]$ 上の一様分布に従うとすると，左辺の n 重積分は $S_{q,n}/S_{p,n}$ $\left(S_{r,n} = \sum_1^n X_k{}^r,\ r = p, q\right)$ の平均値に等しい．同分布の場合の大数の強法則により $n^{-1}S_{r,n} \to E(X_1{}^r) = (r+1)^{-1}$ a.s. であるから $S_{q,n}/S_{p,n} \to (p+1)/(q+1)$ a.s. しかも $0<p<q<1$ により $|S_{q,n}/S_{p,n}| \leqq 1$．有界収束定理により $E(S_{q,n}/S_{p,n}) \to (p+1)/(q+1)$．

§4.6 中心極限定理

$A_1, A_2, \cdots$ が独立事象の列でその確率がすべて p $(0<p<1)$ とする．このとき $A_1, A_2, \cdots, A_n$ の中でおこるものの数を S_n としたとき

$$P\left\{\frac{S_n - np}{\sqrt{np(1-p)}} < s\right\} \longrightarrow \int_{-\infty}^s \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \qquad (n \to \infty)$$

となることが de Moivre, P. S. Laplace によって証明され，**de Moivre-Laplace の定理**として知られている．$X_n = 1_{A_n}$ とすると，$X_1, X_2, \cdots$ は独立確率変数列で，すべて同じ確率法則

$$P\{X_n = 1\} = p, \qquad P\{X_n = 0\} = 1 - p$$

に従っている．この $\{X_n\}$ によって上記の S_n は

$$S_n = \sum_{k=1}^{n} X_k, \qquad n = 1, 2, \cdots$$

であらわされる．しかも

$$E(S_n) = np, \qquad V(S_n) = np(1-p)$$

となることは容易にわかる．このように考えると de Moivre-Laplace の定理は

$T_n = \dfrac{S_n - E(S_n)}{\sqrt{V(S_n)}}$ の確率法則が Gauss 分布 $N(0, 1)$ に近づく

ことを主張している．初等確率論では

$$P\{S_n = k\} = \binom{n}{k} p^k (1-p)^{n-k}$$

に注意し，Stirling の近似式

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n}$$

を用いて，$P\{S_n = k\}$ を評価して de Moivre-Laplace の定理を証明するが，特性関数を利用すれば，つぎのように極めて自然にこの定理が導かれる．

$$T_n = \sum_{k=1}^{n} \frac{X_k - p}{\sqrt{npq}} \qquad (q = 1-p),$$

$$E(e^{izT_n}) = \prod_k E\left(\exp\frac{iz(X_k - p)}{\sqrt{npq}}\right) = (pe^{iz\sqrt{q/np}} + qe^{-iz\sqrt{p/nq}})^n$$

$$= \left\{p\left(1 + iz\sqrt{\frac{q}{np}} - \frac{qz^2}{2np}\right) + q\left(1 - iz\sqrt{\frac{p}{nq}} - \frac{pz^2}{2nq}\right) + o\left(\frac{1}{n}\right)\right\}^n$$

$$= \left(1 - \frac{z^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

$$\longrightarrow e^{-z^2/2} \qquad (n \to \infty),$$

$$(\mathscr{F}P^{T_n})(z) \longrightarrow \mathscr{F}N(z) \qquad (N = N(0, 1)).$$

Glivenko の定理により

$$P^{T_n} \longrightarrow N \qquad (n \to \infty).$$

この証明を見ると，de Moivre-Laplace の定理がもっと一般の場合にも成立することが想像されるであろう．$\{X_n\} \subset L^2$ が独立確率変数列であるとし，

$$S_n = \sum_{k=1}^{n} X_k$$

とおくとき，

$$T_n = \frac{S_n - ES_n}{\sqrt{V(S_n)}}$$

の確率法則が Gauss 分布 $N(0, 1)$ にある条件のもとで近づくことを主張する定理を**中心極限定理**という．特性関数の理論はこの定理の証明のために開発されたといってもよい．$S_n = \sum_1^n X_k$，$E(S_n) = \sum_1^n E(X_k)$ と $\{X_k\}$ の独立性に注意して，

$$\mathscr{F}P^{T_n}(z) = \prod_{k=1}^{n} E\left\{\exp\frac{iz(X_k - EX_k)}{\sqrt{V(S_n)}}\right\}$$

$$= \prod_{k=1}^{n}\left(1 - \frac{V(X_k)}{2V(S_n)}z^2 + \cdots\right)$$

$$\sim \prod_{k=1}^{n} \exp\left\{-\frac{V(X_k)}{2V(S_n)}z^2\right\} = \exp\left\{-\frac{z^2}{2}\right\} \qquad \left(V(S_n) = \sum_1^n V(X_k)\right).$$

これから P^{T_n} が $N(0, 1)$ に近づくことが結論されるであろう．これが中心極限定理を証明する大体の考え方である．

まず準備としてつぎの補題を証明しておこう．

補題 4.2 おのおのの n に対して複素数列

$$\alpha_{n1}, \quad \alpha_{n2}, \quad \cdots, \quad \alpha_{nN} \qquad (N = N(n) < \infty)$$

が与えられ，

$$\alpha_n \equiv \sum_k \alpha_{nk} \longrightarrow \alpha, \qquad \beta_n \equiv \max_k |\alpha_{nk}| \longrightarrow 0 \qquad (n \to \infty),$$

$$\gamma_n \equiv \sum_k |\alpha_{nk}| < b \qquad (b \text{ は } n \text{ に無関係})$$

とする．このとき

$$\prod_k (1 + \alpha_{nk}) \longrightarrow e^{\alpha} \qquad (n \to \infty).$$

証明 これはよく知られた公式

$$e^{\alpha} = \lim_{n \to \infty}\left(1 + \frac{\alpha}{n}\right)^n$$

の一般化である．仮定 $\beta_n \to 0$ により，すべての n, k に対し

$$|\alpha_{nk}| < \frac{1}{2}$$

として一般性を失わない．複素数 z の関数 $\log(1+z)$ は多価であるが，$z=0$ で 0 となる分枝をとれば，$|z|<1$ で1価で

$$\log(1+z) = z-\frac{z^2}{2}+\frac{z^3}{3}-\cdots \qquad (|z|<1)$$

$$= z+\theta z^2 \qquad \left(|z|<\frac{1}{2}\right).$$

ここに θ は絶対値1以下の複素数である．$|\alpha_{nk}|<1/2$ であるから，仮定 $\alpha_n\to\alpha$, $\beta_n\to 0$, $\gamma_n<b$ により

$$\prod_k (1+\alpha_{nk}) = \prod_k \exp\{\log(1+\alpha_{nk})\}$$

$$= \prod_k \exp\{\alpha_{nk}+\theta_{nk}\alpha_{nk}^2\} \qquad (|\theta_{nk}|\leqq 1)$$

$$= \exp\left\{\sum_k \alpha_{nk}+\sum_k \theta_{nk}\alpha_{nk}^2\right\}$$

$$= \exp\left\{\alpha_n+\beta_n\theta_n\sum_k |\alpha_{nk}|\right\} \qquad (|\theta_n|\leqq 1)$$

$$= \exp\{\alpha_n+\beta_n\theta_n' b\} \qquad (|\theta_n'|\leqq 1)$$

$$\longrightarrow e^{\alpha} \qquad (n\to\infty).$$ ∎

補題 4.3　任意の実数 z に対し

$$e^{iz} = \sum_{k=0}^{n}\frac{(iz)^k}{k!}+\frac{\theta|z|^{n+1}}{(n+1)!} \qquad (|\theta|\leqq 1).$$

証明　左辺と右辺の第1項との差を $f(z)$ とすると，

$$f^{(n+1)}(z) = i^{n+1}e^{iz}, \qquad f^{(k)}(0) = 0 \quad (k=0, 1, 2, \cdots, n)$$

であるから

$$f(z) = \int_0^z\int_0^{z_1}\int_0^{z_2}\cdots\int_0^{z_n} i^{n+1}e^{iz_{n+1}}dz_{n+1}dz_n\cdots dz_1,$$

$$|f(z)| \leqq \frac{|z|^{n+1}}{(n+1)!}.$$

これから補題の等式が得られる．∎

中心極限定理がなりたつための条件はいくつも知られているが，つぎに示す Lindeberg の条件は最も一般なものであろう．

定理 4.13 (Lindeberg の中心極限定理)　$\{X_n\}\subset L^2$ は独立で $S_n=\sum_1^n X_k$ とする．すべての $\varepsilon>0$ に対し

$$\frac{1}{V(S_n)}\sum_{k=1}^{n} E((X_k-EX_k)^2, |X_k-EX_k|\geqq\varepsilon\sqrt{V(S_n)}) \longrightarrow 0 \qquad (n\to\infty)$$

ならば，$T_n=(S_n-E(S_n))/\sqrt{V(S_n)}$ の確率法則は，$n\to\infty$ のとき，Gauss 分布 $N(0, 1)$ に近づく．

証明　$Y_{nk}=(X_k-E(X_k))/\sqrt{V(S_n)}$ とおくと

$$E(Y_{nk}) = 0, \qquad V(Y_{nk}) = E(Y_{nk}^2), \qquad T_n = \sum_{k=1}^{n} Y_{nk},$$

$$E(T_n) = 0, \qquad \sum_{k=1}^{n} E(Y_{nk}^2) = \sum_{k=1}^{n} V(Y_{nk}) = V(T_n) = 1$$

となり，定理の仮定は

$$\sum_{k=1}^{n} E(Y_{nk}^2, |Y_{nk}|\geqq\varepsilon) \longrightarrow 0$$

とかける．

さて $\mu_n=P^{T_n}$ とおくと，$\{Y_{nk}, k=1, 2, \cdots, n\}$ の独立性により

$$\mathscr{F}\mu_n = E(e^{izT_n}) = \prod_k E(e^{izY_{nk}}).$$

補題 4.3 により

$$\alpha_{nk} \equiv E(e^{izY_{nk}})-1 = E(e^{izY_{nk}}, |Y_{nk}|<\varepsilon)+E(e^{izY_{nk}}, |Y_{nk}|\geqq\varepsilon)-1$$

$$= E\left(1+izY_{nk}-\frac{1}{2}z^2Y_{nk}^2+\frac{1}{6}\theta_3|zY_{nk}|^3, |Y_{nk}|<\varepsilon\right)$$

$$+E\left(1+izY_{nk}+\frac{1}{2}\theta_2 z^2Y_{nk}^2, |Y_{nk}|\geqq\varepsilon\right)-1 \qquad (|\theta_2|, |\theta_3|\leqq 1)$$

$$= izE(Y_{nk})-\frac{1}{2}z^2E(Y_{nk}^2)+\frac{1}{2}z^2E(Y_{nk}^2, |Y_{nk}|\geqq\varepsilon)$$

$$+\frac{1}{6}|z|^3E(\theta_3|Y_{nk}|^3, |Y_{nk}|<\varepsilon)+\frac{1}{2}z^2E(\theta_2Y_{nk}^2, |Y_{nk}|\geqq\varepsilon)$$

$$= -\frac{1}{2}z^2E(Y_{nk}^2)+\frac{1}{2}z^2(1+\theta_2')E(Y_{nk}^2, |Y_{nk}|\geqq\varepsilon)$$

$$+\frac{1}{6}|z|^3\theta_3'\varepsilon E(Y_{nk}^2) \qquad (|\theta_2'|, |\theta_3'|\leqq 1).$$

$$\max_k E(Y_{nk}^2, |Y_{nk}|\geqq\varepsilon) \leqq \sum_k E(Y_{nk}^2, |Y_{nk}|\geqq\varepsilon) \longrightarrow 0 \qquad (n\to\infty),$$

$$\max_k E(Y_{nk}^2) \leqq \varepsilon^2+\sum_k E(Y_{nk}^2, |Y_{nk}|\geqq\varepsilon) \longrightarrow \varepsilon^2 \qquad (n\to\infty)$$

により

$$\limsup_{n\to\infty} \max_k |\alpha_{nk}| \leqq \frac{1}{2}z^2\varepsilon^2+\frac{1}{6}|z|^3\varepsilon^3 \longrightarrow 0 \qquad (\varepsilon \downarrow 0).$$

したがって

$$\max_k |\alpha_{nk}| \longrightarrow 0 \qquad (n\to\infty).$$

つぎに

$$\sum_k \alpha_{nk} = -\frac{1}{2}z^2+\frac{1}{2}z^2(1+\theta_2'')\sum_k E(Y_{nk}^2, |Y_{nk}|\geqq\varepsilon)$$
$$+\frac{1}{6}|z|^3\theta_3''\varepsilon \qquad (|\theta_2''|, |\theta_3''|\leqq 1).$$

これから上と同じ論法で

$$\sum_k \alpha_{nk} \longrightarrow -\frac{1}{2}z^2 \qquad (n\to\infty).$$

最後に

$$\sum_k |\alpha_{nk}| \leqq \frac{1}{2}z^2+z^2+\frac{1}{6}|z|^3\varepsilon.$$

これから補題4.2により

$$\mathscr{F}\mu_n(z) = \prod_k E(e^{izY_{nk}}) = \prod_n (1+\alpha_{nk}) \longrightarrow e^{-z^2/2}$$

が得られ，Glivenko の定理により，$\mu_n\to N(0,1)$ を得る．∎

上の Lindeberg の条件は暗黙のうちに

$$v\equiv \lim_{n\to\infty} V(S_n) > 0$$

を含んでいることはいうまでもないが，さらに $v=\infty$ を含んでいることを証明しよう．もし $0<v<\infty$ であれば，ある l に対し，$V(X_l)>0$，したがって

$$v_{l,\varepsilon} \equiv E((X_l-EX_l)^2, |X_l-EX_l|\geqq\varepsilon\sqrt{V(S_n)})$$
$$\geqq E((X_l-EX_l)^2, |X_l-EX_l|\geqq\varepsilon\sqrt{v}) \longrightarrow V(X_l)>0 \qquad (\varepsilon\downarrow 0).$$

それゆえ十分小さい $\varepsilon>0$ に対し $v_{l,\varepsilon}>0$ となり，Lindeberg の条件の左辺は $n>l$ のとき $v_{l,\varepsilon}/v$ より大きく，0 に収束することは不可能である．したがって Lindeberg の条件は

$$v=\lim_{n\to\infty} V(S_n)=\infty \quad すなわち \quad \sum_{n=1}^{\infty} V(X_n)=\infty$$

を含んでいる．この条件のもとでは，定理4.10により

$$T_n^{\alpha} = \frac{S_n-E(S_n)}{V(S_n)^{\alpha}} \longrightarrow 0 \text{ a.s.}$$

が任意の $\alpha>1/2$ に対してなりたつから，T_n^{α} の確率法則は δ に収束する．$\alpha=1/2$ とおくと，$T_n^{1/2}=T_n$ となり，Lindeberg の定理により，その確率法則は $N(0,1)$ に収束する．

$\alpha>1/2$ のときには T_n^{α} は 0 に概収束するが，$T_n^{1/2}\equiv T_n$ はある確率変数に概収束するであろうか．もしこういうことがおこれば，$m,n\to\infty$ のとき

$$T_n-T_m \longrightarrow 0 \text{ a.s.}$$

さて $m=m(n)$ を n に対しはるかに大きくとって

$$\frac{V(S_n)}{V(S_m)} \longrightarrow 0 \qquad (n\to\infty)$$

となるようにしておこう．さて

$$T_n-T_m = \frac{\sum_1^n (X_k-E(X_k))}{\sqrt{V(S_n)}} - \frac{\sum_1^m (X_k-E(X_k))}{\sqrt{V(S_m)}}$$
$$= \left(1-\sqrt{\frac{V(S_n)}{V(S_m)}}\right)T_n - \frac{1}{\sqrt{V(S_m)}}\sum_{n+1}^{m}(X_k-E(X_k))$$
$$\equiv U_n-V_n.$$

T_n の確率法則は $N(0,1)$ に近づくから，U_n の確率法則も同様である．また $\{X_1, X_2, \cdots, X_m\}$ は独立であるから，U_n と V_n とは独立である．したがって特性関数の乗法性により

$$|E(e^{iz(T_n-T_m)})| = |E(e^{izU_n})||E(e^{-izV_n})|$$
$$\leqq |E(e^{izU_n})| \longrightarrow e^{-z^2/2} \qquad (n\to\infty).$$

$n\to\infty$ のとき，$T_n-T_m\,(m=m(n))$ は 0 に概収束するから，左辺は 1 に収束し，上の式から

$$1\leqq e^{-z^2/2} \qquad (矛盾!)$$

に到達する．かくしてつぎの定理が得られる．

定理 4.14　Lindeberg の条件のもとで $T_n=(S_n-E(S_n))/\sqrt{V(S_n)}$ は概収束しない．(実は確率収束もしないことが上の証明からわかる)．

例題 4.6　(i), (ii), (iii) では $S_n=\sum_1^n X_k$ とする．

(i)　$M\equiv\sup_n \|X_n\|_\infty<\infty$，$\sum_n V(X_n)=\infty$ をみたす独立な $\{X_n\}\subset L^2$ に対して

は中心極限定理がなりたつことを証明せよ．

［ヒント］ 仮定により $|E(X_n)|\leqq M$, したがって $|X_n-E(X_n)|\leqq 2M$ a.s. 任意の $\varepsilon>0$ に対し，n を十分大きくとれば

$$E((X_k-E(X_k))^2, |X_k-E(X_k)|\geqq \varepsilon V(S_n)) = 0, \qquad k=1,2,\cdots,n.$$

(ii) $\{X_n\}\subset L^2$ が独立で，正の分散をもつ同じ分布 μ に従うときには，中心極限定理がなりたつことを示せ．

［ヒント］ Lindeberg の条件の式の左辺は

$$\frac{1}{V(\mu)}\int_{|x-M(\mu)|>\varepsilon\sqrt{nV(\mu)}}(x-M(\mu))^2\mu(dx)$$

に等しく，これは，$n\to\infty$ のとき，0に近づく．

(iii) 独立な $\{X_n\}\subset L^2$ が Ljapunov の条件

$$\frac{1}{\sqrt{V(S_n)}^{(2+\delta)}}\sum_{k=1}^{n}E(|X_k-E(X_k)|^{2+\delta}) \longrightarrow 0 \qquad (\delta \text{ はある正定数})$$

をみたすならば，中心極限定理がなりたつことを示せ．

［ヒント］ $E(|X_k-E(X_k)|^{2+\delta})$

$$\geqq E(|X_k-E(X_k)|^{2+\delta}, |X_k-E(X_k)|\geqq\varepsilon\sqrt{V(S_n)})$$

$$\geqq \varepsilon^\delta\sqrt{V(S_n)}^\delta E(|X_k-E(X_k)|^2, |X_k-E(X_k)|\geqq\varepsilon\sqrt{V(S_n)})$$

に注意して Lindeberg の条件を検証せよ．

(iv) 下の等式を証明せよ．

$$\lim_{n\to\infty}e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\frac{1}{2}.$$

［ヒント］ $X_1, X_2, \cdots$ が独立で，すべて同じ Poisson 分布 p_1 (§2.5) に従うとする．$E(X_k)=V(X_k)=1$ であるから $E(S_n)=n$, $V(S_n)=n$ となる．問題(ii)により中心極限定理がなりたつから $T_n=(S_n-n)/\sqrt{n}$ の確率法則は $N\equiv N(0,1)$ に近づく．N は連続分布であるから，$P\{S_n\leqq n\}=P\{T_n\leqq 0\}\to 1/2$ がでる．S_n は Poisson 分布 p_n に従うことに注意すれば上の等式が導かれる．

§4.7 重複対数の法則

$\{X_n\}\subset L^2$ を独立とし

$$S_n=\sum_{k=1}^{n}X_k$$

とおく．$V(S_n)\to\infty$ と仮定すると，定理4.10により

$$\lim_{n\to\infty}\frac{S_n-E(S_n)}{V(S_n)^{1/2+\varepsilon}}=0 \text{ a.s.}$$

これは $n\to\infty$ のとき $|S_n-E(S_n)|$ が $V(S_n)^{1/2+\varepsilon}$ に比して極めて小さいことを意味している．ここで ε は正数であればいかに小さくともよい．しかし $\varepsilon=0$ とおくと，定理4.14により $(S_n-E(S_n))/V(S_n)^{1/2}$ は概収束しない．さらにつぎの定理がなりたつ．

定理 4.15 Lindeberg の条件のもとでは，

$$\limsup_{n\to\infty}\frac{S_n-E(S_n)}{V(S_n)^{1/2}}=\infty \text{ a.s.},$$

$$\liminf_{n\to\infty}\frac{S_n-E(S_n)}{V(S_n)^{1/2}}=-\infty \text{ a.s.}$$

証明 $E(X_n)\equiv 0$ (したがって $E(S_n)\equiv 0$) として一般性を失わない．$\{X_1, X_2, \cdots\}$ が Lindeberg の条件をみたすから，任意の n に対し，$\{X_n, X_{n+1}, \cdots\}$ も同じ条件をみたし，中心極限定理により

$$\lim_{m\to\infty}P\left\{\frac{S_m-S_n}{\sqrt{V(S_m-S_n)}}>\alpha\right\}=N(\alpha,\infty)>0 \qquad (\alpha>0,\ N=N(0,1)).$$

したがって急激に増加する自然数列 $r(1)<r(2)<\cdots$ を適当にとって，

$$\frac{\sqrt{V(S_{r(n)})}}{V(S_{r(n-1)})}\longrightarrow\infty, \qquad \beta=\inf_n P\left\{\frac{S_{r(n)}-S_{r(n-1)}}{\sqrt{V(S_{r(n)}-S_{r(n-1)})}}>\alpha\right\}>0$$

とすることができる．$\{X_n\}$ の独立性により，$\{S_{r(n)}-S_{r(n-1)}, n=1,2,\cdots\}$ が独立，したがって Borel の定理によりほとんど確実に

$$S_{r(n)}-S_{r(n-1)}>\alpha\sqrt{V(S_{r(n)}-S_{r(n-1)})} \text{ i.o.},$$

したがって

$$\frac{S_{r(n)}}{\sqrt{V(S_{r(n)})}}>\frac{V(S_{r(n-1)})}{\sqrt{V(S_{r(n)})}}\frac{S_{r(n-1)}}{V(S_{r(n-1)})}+\alpha\sqrt{\frac{V(S_{r(n)}-S_{r(n-1)})}{V(S_{r(n)})}} \text{ i.o.}$$

$n\to\infty$ のとき，右辺の第1項は0に近づき，第2項の平方根の中の式は，

$$V(S_{r(n)}-S_{r(n-1)})=V(S_{r(n)})-V(S_{r(n-1)})$$

により，1に近づく．ゆえに上の式から

$$\limsup_{n\to\infty}\frac{S_{r(n)}}{\sqrt{V(S_{r(n)})}}\geqq\alpha \text{ a.s.},$$

$$\limsup_{n\to\infty}\frac{S_n}{\sqrt{V(S_n)}}\geqq\alpha\ \text{a.s.}$$

$\alpha\uparrow\infty$ として

$$\limsup_{n\to\infty}\frac{S_n}{\sqrt{V(S_n)}}=\infty\ \text{a.s.}$$

を得る．これで定理の第1式は証明された．ここで $\{X_n\}$ のかわりに $\{-X_n\}$ を考えて，第2式が得られる．▌

上の定理により $S_n-E(S_n)$ は無限回 $\sqrt{V(S_n)}$ より大きく，また $-\sqrt{V(S_n)}$ より小さくなるから，結局 $S_n-E(S_n)$ の概発散の型は **'振幅無限増大の振動'** である．このことから，とくに $P\{X_n=-1\}=P\{X_n=1\}=1/2$ のときには，ほとんど確実に無限回 $S_n=0$ となることがわかる．

Lindeberg の条件よりもさらに仮定を強めると

$$\limsup_{n\to\infty}\frac{S_n-E(S_n)}{(2V(S_n)\log^2 V(S_n))^{1/2}}=1\ \text{a.s.},$$

$$\liminf_{n\to\infty}\frac{S_n-E(S_n)}{(2V(S_n)\log^2 V(S_n))^{1/2}}=-1\ \text{a.s.},$$

$$\limsup_{n\to\infty}\frac{|S_n-E(S_n)|}{(2V(S_n)\log^2 V(S_n))^{1/2}}=1\ \text{a.s.}$$

がなりたつことが示される．これは定理 4.10, 4.15 よりもはるかに精密な結果であって，この型の定理を**重複対数の法則**とよぶ．この著しい結果は $\{X_n\}$ が適当な条件のもとで A. Hinčin によって始めて証明され，その後次第に一般化され，また精密化された．ここでは定理の証明の基本的な考え方を示すために，$\{X_n\}$ がすべて Gauss 分布に従うとき（このとき X_n を **Gauss 変数**という）を考察する．

定理 4.16　$\{X_n\}$ が独立な Gauss 変数列で，$V(S_n)\to\infty$, $V(X_n)/V(S_n)\to 0$ のときに重複対数の法則がなりたつ．

証明　X_n のかわりに $X_n-E(X_n)$ を考えることにより，$E(X_n)=0$（したがって $E(S_n)=0$）として一般性を失わない．まず

$$\text{(i)}\quad \limsup_{n\to\infty}\frac{S_n}{(2V(S_n)\log^2 V(S_n))^{1/2}}\leqq 1\ \text{a.s.}$$

を証明しよう．任意の $\varepsilon>0$ に対し自然数列

$$n(1)<n(2)<n(3)<\cdots$$

を帰納的に

$$n(1)=1,\qquad n(k)=\min\{n\mid V(S_n)\geqq(1+\varepsilon)V(S_{n(k-1)})\}$$

によって定めると，

$$V(S_{n(k)})\geqq(1+\varepsilon)V(S_{n(k-1)})>V(S_{n(k)-1}).$$

仮定 $V(X_n)/V(S_n)\to 0$ により，$V(S_n)/V(S_{n-1})\to 1$．したがって十分大きい k に対しては

$$V(S_{n(k)})<(1+\varepsilon)V(S_{n(k)-1})<(1+\varepsilon)^2V(S_{n(k-1)}).$$

これからすべての k に対して

$$c_1(1+\varepsilon)^k\leqq V(S_{n(k)})<c_1(1+\varepsilon)^{2k}.$$

ここで c_1 は k に関係しない定数である．（以下 $c_2, c_3, \cdots$ はすべて k に関係しない定数とする．）

さて

$$T_k=\max\{S_j\mid 1\leqq j\leqq n(k)\}$$

とおいて

$$p_k=p_k(\varepsilon,\delta)=P\{T_k\geqq(2(1+\delta)V(S_{n(k)})\log^2 V(S_{n(k)}))^{1/2}\}\qquad(\delta>0)$$

を評価しよう．$X_1, X_2, \cdots$ が独立で平均値0の Gauss 分布に従うから，S_n-S_j $(n\geqq j)$ は平均値0の Gauss 分布に従い

$$P\{S_n-S_j\geqq 0\}\geqq\frac{1}{2}\qquad(n>j\text{ のときには等号成立}).$$

任意の $a>0$ に対し

$$\begin{aligned}
&P\left\{\max_{j=1}^{n}S_j\geqq a\right\}\\
&\quad=\sum_{k=1}^{n}P\left\{\max_{j=1}^{k-1}S_j<a\leqq S_k\right\}\qquad\left(k=1\text{ のときには }\max_{j=1}^{k-1}S_j=0\text{ とする}\right),\\
&P\{S_n\geqq a\}\\
&\quad\geqq\sum_{k=1}^{n}P\left\{\max_{j=1}^{k-1}S_j<a\leqq S_k,\ S_n-S_k\geqq 0\right\}\\
&\quad=\sum_{k=1}^{n}P\left\{\max_{j=1}^{k-1}S_j<a\leqq S_k\right\}P\{S_n-S_k\geqq 0\}\qquad(\{X_k\}\text{ の独立性により})\\
&\quad\geqq\frac{1}{2}\sum_{k=1}^{n}P\left\{\max_{j=1}^{k-1}S_j<a\leqq S_k\right\}.
\end{aligned}$$

ゆえに

$$P\left\{\max_{j=1}^{n} S_j \geqq a\right\} \leqq 2P\{S_n \geqq a\},$$

a のかわりに $a\sqrt{V(S_n)}$ とおいて

$$P\left\{\max_{j=1}^{n} S_j \geqq a\sqrt{V(S_n)}\right\} \leqq 2P\{S_n \geqq a\sqrt{V(S_n)}\} = 2\int_a^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx,$$

$$\int_a^\infty e^{-x^2/2} dx = \int_a^\infty \frac{1}{x} e^{-x^2/2} x dx = -\frac{1}{x} e^{-x^2/2}\Big|_a^\infty - \int_a^\infty \frac{1}{x^2} e^{-x^2/2} dx$$

$$\leqq \frac{1}{a} e^{-a^2/2} \leqq e^{-a^2/2} \qquad (a>1),$$

$$P\left\{\max_{j=1}^{n} S_j \geqq a\sqrt{V(S_n)}\right\} \leqq e^{-a^2/2} \qquad (a>1).$$

$n=n(k)$, $a=(2(1+\delta)\log^2 V(S_{n(k)}))^{1/2}$ とおいて

$$\begin{aligned} p_k &\equiv P\{T_k \geqq (2(1+\delta) V(S_{n(k)}) \log^2 V(S_{n(k)}))^{1/2}\} \\ &\leqq \exp\{-(1+\delta)\log^2 V(S_{n(k)})\} \\ &= (\log V(S_{n(k)}))^{-(1+\delta)} \\ &\leqq (\log(c_1(1+\varepsilon)^k))^{-(1+\delta)} \\ &\leqq c_2 k^{-(1+\delta)}. \end{aligned}$$

ゆえに

$$\sum_k p_k < \infty.$$

Borel-Cantelli の補題により

$$P\{T_k < (2(1+\delta) V(S_{n(k)}) \log^2 V(S_{n(k)}))^{1/2} \text{ f.e.}\} = 1.$$

すなわち

$$\text{(ii)} \quad \limsup_{k\to\infty} \frac{T_k}{(2V(S_{n(k)}) \log^2 V(S_{n(k)}))^{1/2}} \leqq (1+\delta)^{1/2} \text{ a.s.}$$

さて

$$n(k-1) \leqq n < n(k)$$

となるすべての n に対して

$$\frac{S_n}{(2V(S_n)\log^2 V(S_n))^{1/2}} \leqq \frac{T_k}{(2V(S_{n(k-1)})\log^2 V(S_{n(k-1)}))^{1/2}}$$

$$= \frac{T_k}{(2V(S_{n(k)})\log^2 V(S_{n(k)}))^{1/2}} \left[\frac{V(S_{n(k)})}{V(S_{n(k-1)})} \cdot \frac{\log^2 V(S_{n(k)})}{\log^2 V(S_{n(k-1)})}\right]^{1/2}$$

$x>e^e$ では $\log^2 x/x$ は減少関数であるから,

$$e^e < x < y \Longrightarrow \frac{\log^2 y}{y} < \frac{\log^2 x}{x} \Longrightarrow \frac{\log^2 y}{\log^2 x} < \frac{y}{x}.$$

ゆえに上の式の $[\ \ ]^{1/2}$ は $n(k)$ が十分大きいときには

$$\frac{V(S_{n(k)})}{V(S_{n(k-1)})} \quad (<(1+\varepsilon)^2)$$

より小さい. したがって (ii) により

$$\limsup_{n\to\infty} \frac{S_n}{(2V(S_n)\log^2 V(S_n))^{1/2}} \leqq (1+\delta)^{1/2}(1+\varepsilon)^2 \text{ a.s.}$$

$\varepsilon\downarrow 0$, $\delta\downarrow 0$ として (i) が得られる. ここで $\{X_n\}$ のかわりに $\{-X_n\}$ を考えると,

$$\liminf_{n\to\infty} \frac{S_n}{(2V(S_n)\log^2 V(S_n))^{1/2}} \geqq -1 \text{ a.s.}$$

が得られ, これと (i) と合せて

$$\text{(iii)} \quad \limsup_{n\to\infty} \frac{|S_n|}{(2V(S_n)\log^2 V(S_n))^{1/2}} \leqq 1 \text{ a.s.}$$

が得られる.

つぎに

$$\text{(iv)} \quad \limsup_{n\to\infty} \frac{S_n}{(2V(S_n)\log^2 V(S_n))^{1/2}} \geqq 1 \text{ a.s.}$$

を示そう. これがいえると, 重複対数の法則の第1式が得られ, 再び $\{X_n\}$ のかわりに $\{-X_n\}$ を考えて第2式が得られ, さらに両式を合せて第3式が得られて, 定理の証明が完了する.

$\alpha>1$ に対し自然数列

$$1 = m(1) < m(2) < m(3) < \cdots$$

を

$$V(S_{m(k)}) \geqq \alpha V(S_{m(k-1)})$$

となるようにとると,

$$V(S_{m(k)}) > c_3 \alpha^k$$

となる. $\{X_n\}$ が独立な Gauss 変数列であるから,

$$\Delta_k = S_{m(k)} - S_{m(k-1)}, \qquad k = 1, 2, \cdots$$

も同様であることがわかる.

さて $\delta>0$ に対し

$$q_k = q_k(\alpha, \delta) = P\{\Delta_k > (2(1-\delta)\, V(\Delta_k) \log^2 V(\Delta_k))^{1/2}\}$$

を評価しよう.

$$P\{\Delta_k > a(V(\Delta_k))^{1/2}\} = \int_a^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx,$$

部分積分を2回行って

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{a}e^{-a^2/2} - \frac{1}{a^3}e^{-a^2/2} + \int_a^\infty \frac{3}{x^4} e^{-x^2/2} dx\right)$$

$$\geqq \frac{1}{4a}e^{-a^2/2} \qquad (a>2).$$

$a = (2(1-\delta)\log^2 V(\Delta_k))^{1/2}$ とおいて, 十分大きい k に対しては

$$q_k \geqq c_4 (\log^2 V(\Delta_k))^{-1/2} (\log V(\Delta_k))^{-1+\delta}$$

$$\geqq c_5 (\log V(\Delta_k))^{-1} \qquad \left(\because \lim_{x\to\infty} (\log^2 x)^{-1/2} (\log x)^{\delta} = \infty\right),$$

$$V(\Delta_k) = V(S_{m(k)}) - V(S_{m(k-1)}) \geqq V(S_{m(k)})\left(1-\frac{1}{\alpha}\right)$$

$$\geqq c_3 \alpha^k \left(1 - \frac{1}{\alpha}\right)$$

に注意して

$$q_k \geqq c_6 k^{-1} \qquad (k \text{ が十分大きいとき}),$$

ゆえに

$$\sum_k q_k = \infty.$$

$\{\Delta_n\}$ が独立であるから, Borel の定理により

$$P\{\Delta_k > (2(1-\delta)\, V(\Delta_k) \log^2 V(\Delta_k))^{1/2} \text{ i.o.}\} = 1,$$

したがって

$$\text{(v)} \quad \limsup_{k\to\infty} \frac{\Delta_k}{(2V(\Delta_k)\log^2 V(\Delta_k))^{1/2}} \geqq (1-\delta)^{1/2} \text{ a.s.}$$

$$\frac{S_{m(k)}}{(2V(S_{m(k)})\log^2 V(S_{m(k)}))^{1/2}} \geqq \frac{\Delta_k}{(2V(S_{m(k)})\log^2 V(S_{m(k)}))^{1/2}} - \frac{|S_{m(k-1)}|}{(2V(S_{m(k)})\log^2 V(S_{m(k)}))^{1/2}},$$

$$V(\Delta_k),\ V(S_{m(k-1)}) \leqq V(S_{m(k)}),$$

$$\frac{V(\Delta_k)}{V(S_{m(k)})} \geqq 1 - \frac{1}{\alpha}, \qquad \frac{V(S_{m(k-1)})}{V(S_{m(k)})} \leqq \frac{1}{\alpha},$$

$$e^e < x < y \Longrightarrow 1 < \frac{\log^2 y}{\log^2 x} < \frac{y}{x}$$

を考慮して, 十分大きい k に対し

$$\frac{S_{m(k)}}{(2V(S_{m(k)})\log^2 V(S_{m(k)}))^{1/2}}$$

$$\geqq \frac{\Delta_k}{(2V(\Delta_k)\log^2 V(\Delta_k))^{1/2}}\left(1-\frac{1}{\alpha}\right) - \frac{|S_{m(k-1)}|}{(2V(S_{m(k-1)})\log^2 V(S_{m(k-1)}))^{1/2}}\left(\frac{1}{\alpha}\right)^{1/2}.$$

したがって (iii), (v) により

$$\limsup_{k\to\infty} \frac{S_{m(k)}}{(2V(S_{m(k)})\log^2 V(S_{m(k)}))^{1/2}} \geqq (1-\delta)^{1/2}\left(1-\frac{1}{\alpha}\right) - \left(\frac{1}{\alpha}\right)^{1/2} \text{a.s.}$$

ゆえに

$$\limsup_{n\to\infty} \frac{S_n}{(2V(S_n)\log^2 V(S_n))^{1/2}} \geqq (1-\delta)^{1/2}\left(1-\frac{1}{\alpha}\right) - \left(\frac{1}{\alpha}\right)^{1/2} \text{a.s.}$$

$\delta\downarrow 0$, $\alpha\uparrow\infty$ として (iv) を得る. ▌

例題 4.7 定理 4.16 の (i) 式は, 任意の $\delta>0$ に対し

$$P\{S_n - E(S_n) < (1+\delta)(2V(S_n)\log^2 V(S_n))^{1/2} \text{ f.e.}\} = 1,$$

$$P\{S_n - E(S_n) > (1-\delta)(2V(S_n)\log^2 V(S_n))^{1/2} \text{ i.o.}\} = 1$$

を示している. 証明の鍵となるのは, U の確率法則が $N(0, V)$ のとき

$$\frac{1}{4a}e^{-a^2/2} \leqq P(U > a\sqrt{V}) \leqq e^{-a^2/2} \qquad (\alpha>2)$$

となることを利用して,

$$\sum_k p_k < \infty, \qquad \sum_k q_k = \infty$$

を証明して Borel-Cantelli の補題または Borel の定理にもちこむという論法である. この点に注意して定理 4.16 の (i) 式 ((ii) 式, (iii) 式も同様) がつぎのように精密化できることを証明せよ. 任意の $\delta>0$ に対し

$$P\{S_n-E(S_n)<(2V(S_n)(\log^2 V(S_n)+(1+\delta)\log^3 V(S_n)))^{1/2}\ \text{f.e.}\}=1,$$
$$P\{S_n-E(S_n)>(2V(S_n)(\log^2 V(S_n)+(1-\delta)\log^3 V(S_n)))^{1/2}\ \text{i.o.}\}=1.$$

§4.8 Gauss の誤差論

実験である量を測定すると誤差があることは誰でも経験することである．誤差には系統誤差 (systematic error) と偶然誤差 (random error) とある．前者は原因が単純で注意深くすれば避けることができる．例えば，天秤で物の重さを計るとき支点から皿の中心への距離が左右わずかに違うことからおこる誤差である．この距離を d_1, d_2 とし，真の重さ w のものを左の皿において，分銅を右の皿におけば，測定値 w_1 は wd_1/d_2 となり，入れかえて計れば，測定値 w_2 は wd_2/d_1 となるはずであるから，

$$w_1w_2=w^2 \quad すなわち \quad w=\sqrt{w_1w_2}$$

となる．すなわち，このときには2測定値の幾何平均をとって真の値が正確に得られる．

偶然誤差は実験室の中の空気の流れ，床の微動，重力の微小な変化など制御できない無数の微小な原因の集積としておこる確率変数である．誤差論は偶然誤差を除去する方法として Gauss の考案したものである．誤差論の出発点は偶然誤差を平均値0の Gauss 分布に従う確率変数と考えることである．

偶然誤差 X が Gauss 分布に従うことの説明としては，それが無数の独立な微小偶然誤差 $X_1, X_2, \cdots, X_n$ の和と考えて X の確率法則が Gauss 分布に近いことを証明するという考えが最も自然であろう．数学的にはつぎの定理を証明すればよい．

定理 4.17 おのおのの n に対して

$$X_{n1},\quad X_{n2},\quad \cdots,\quad X_{nN} \qquad (N=N(n))$$

が独立とし，$S_n=\sum_{k=1}^{N(n)} X_{nk}$ とする．

$$\varepsilon_n=\sup_{k,\omega}|X_{nk}(\omega)|\longrightarrow 0,$$
$$E(S_n)\equiv\sum_k E(X_{nk})\longrightarrow m,$$
$$V(S_n)\equiv\sum_k V(X_{nk})\longrightarrow v$$

ならば，S_n の確率法則 μ_n は $N(m,v)$ に近づく．

証明 $\theta_{nk}, \theta_{nk}', \cdots$ はすべて絶対値1以下の複素数をあらわす．$\{X_{nk}, k=1,2,\cdots,N\}$ の独立性により，

$$\mathscr{F}\mu_n(z)=E(e^{izS_n})=\prod_k E(e^{izX_{nk}}).$$

$m_{nk}=E(X_{nk})$，$m_n=E(S_n)$ とおくと，$m_n=\sum_k m_{nk}$ であるから，中心極限定理の証明と同様に

$$\begin{aligned}
右辺 &= e^{izm_n}\prod_k E(e^{iz(X_{nk}-m_{nk})})\\
&= e^{izm_n}\prod_k\left(1-\frac{z^2}{2}V(X_{nk})+\frac{z^3}{6}E(|X_{nk}-m_{nk}|^3)\theta_{nk}\right)\\
&= e^{izm_n}\prod_k\left(1-\frac{z^2}{2}V(X_{nk})+\frac{z^3}{6}V(X_{nk})(2\varepsilon_n)\theta_{nk}'\right)\\
&\qquad (\because\ |X_{nk}-m_{nk}|^3\leqq (X_{nk}-m_{nk})^2(|X_{nk}|+|m_{nk}|))\\
&\longrightarrow \exp\left(izm-\frac{z^2}{2}v\right).
\end{aligned}$$
∎

偶然誤差の確率法則は Gauss 分布とみなされるが，普通その平均値 m は0とする．（もし $m\neq 0$ のときには，系統誤差が除かれていないためで，実験方法を検討して見る必要があろう.）したがって偶然誤差の確率法則は $N(0,v)$ の型である．

系統誤差を除去したとすれば，測定値 X は真の値 a と偶然誤差の和になるから，X の確率法則は $N(a,v)$ となる．

$$P\{|X-a|>t\sqrt{v}\}=2\int_t^\infty \frac{1}{\sqrt{2\pi}}e^{-s^2/2}ds=\sqrt{\frac{2}{\pi}}\int_t^\infty e^{-s^2/2}ds$$

は a, v に無関係に t だけで定まるから，$\sqrt{v}$ は誤差の大きさを調べる目安となる量である．$\sqrt{v}$ は $\sqrt{V(X-a)}$ すなわち $\sigma(X-a)$ (誤差 $X-a$ の標準偏差) である．標準偏差の名はこの事実を背景にして生れたのである．v が小さい程測定精度がよくなる．

測定の精度をあげるために，何回も測定して平均をとるということが行われる．この数学的意味はつぎのようである．n 回測定するとし，第 k 回目の測定値を X_k とする．n 回の測定は全く無関係におこなうとすれば，$X_1, X_2, \cdots, X_n$ は独立な確率変数と見てよい．また測定方法は同じとすれば，$X_1, X_2, \cdots, X_n$ はすべて

同じ分布 $N(m,v)$ に従うと考えられる．さて $X_1, X_2, \cdots, X_n$ の平均を $\bar{X}$ とする．すなわち

$$\bar{X} = \frac{1}{n}(X_1+X_2+\cdots+X_n).$$

ここで注意すべきは $\bar{X}$ も確率変数であるということである．

$$E(e^{iz\bar{X}}) = \prod_{k=1}^{n} E(e^{izX_k/n})$$
$$= \prod_{k=1}^{n} \exp\left\{\frac{iaz}{n}-\frac{vz^2}{2n^2}\right\}$$
$$= \exp\left\{iaz-\frac{v}{2n}z^2\right\}$$

により，$\bar{X}$ の確率法則は $N(a, v/n)$ に等しい．n が大きくなると v/n は小さくなるから，測定の精度があがるのである．

上の事実は偶然誤差が Gauss 分布に従うとしたことから出たので，もし中心値 a の Cauchy 分布に従うとし，

$$P\{X_k \in E\} = \frac{c}{\pi}\int_E \frac{dx}{c^2+(x-a)^2}, \qquad k=1,2,\cdots,n$$

として見ると，

$$E(e^{izX_k}) = e^{iaz-c|z|},$$
$$E(e^{iz\bar{X}}) = \left(\exp\left\{\frac{iaz}{n}-\frac{c|z|}{n}\right\}\right)^n = e^{iaz-c|z|},$$
$$P^{\bar{X}} = P^{X_k}, \qquad k=1,2,\cdots,n$$

となる．これは多数回の測定値を平均しても，1回だけの測定と同じことであることを示している．Gauss 分布，Cauchy 分布の密度のグラフは一見よく似ているが，本質的な差があることがわかる．

もとに戻って平均値 $\bar{X}$ をとるのに重み $\alpha_1, \alpha_2, \cdots, \alpha_n$ をつけて，これを $\bar{X}_\alpha$ であらわそう．

$$\alpha_1, \alpha_2, \cdots, \alpha_n \geqq 0, \qquad \sum_{k=1}^{n}\alpha_k = 1$$

としておいてよい．こうすれば

$$\bar{X}_\alpha = \sum_k \alpha_k X_k$$

となり，その確率法則は $N(m, v_\alpha)$ $(v_\alpha = \sum_k \alpha_k^2 v)$ である．$\sum_k \alpha_k = 1$ により v_α は

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = \frac{1}{n}$$

のとき最小となるから，同じ重みによる平均値 $\bar{X}$ が最良のものであることがわかる．

例題 4.8 三角形の内角 $\theta_1, \theta_2, \theta_3$ を同じ測定法で無関係に計って X_1, X_2, X_3 を得た．$\theta_1, \theta_2, \theta_3$ を評価する最良の式 (X_1, X_2, X_3 の1次式) を求めよ．

［ヒント］ 求める最良の式を

$$\bar{X}_i = a_i + \sum_{k=1}^{3} a_{ik}X_k, \qquad i=1,2,3$$

とする．仮定により，X_1, X_2, X_3 は独立で，その確率法則は $N(\theta_1, v), N(\theta_2, v), N(\theta_3, v)$ となる．(v が等しいのは測定法が同じであるから．) したがって $\bar{X}_1$ の確率法則は

$$N(\bar{\theta}_1, v_1) \qquad \left(\bar{\theta}_1 = a_1 + \sum_k a_{1k}\theta_1,\ \ v_1 = \sum_k a_{1k}^2 v\right)$$

である．θ_1 は真の値であるから，

$$\bar{\theta}_1 = \theta_1 \quad \text{すなわち} \quad a_1 + \sum_k a_{1k}\theta_k = \theta_1.$$

これは，三角形の内角の条件

$$\theta_1+\theta_2+\theta_3 = \pi$$

がみたされる限り，常に成立しなければならない．したがって

$$\frac{a_{11}-1}{1} = \frac{a_{12}}{1} = \frac{a_{13}}{1} = \frac{-a_1}{\pi}.$$

この条件のもとでは

$$v_1/v = \sum_k a_{1k}^2 = a_{11}^2 + 2(a_{11}-1)^2 = 3a_{11}^2 - 4a_{11} + 2$$

となるが，これを最小にする a_{11} の値は 2/3．したがって

$$a_{12} = -\frac{1}{3}, \qquad a_{13} = -\frac{1}{3}, \qquad a_1 = \frac{\pi}{3}$$

となるから，結局

$$\bar{X}_1 = \frac{\pi}{3} + \frac{2}{3}X_1 - \frac{1}{3}X_2 - \frac{1}{3}X_3$$

が θ_1 の最良の評価式である．同様に θ_2, θ_3 の最良評価式は

$$\bar{X}_2 = \frac{\pi}{3} - \frac{1}{3}X_1 + \frac{2}{3}X_2 - \frac{1}{3}X_3,$$

$$\bar{X}_3 = \frac{\pi}{3} - \frac{1}{3}X_1 - \frac{1}{3}X_2 + \frac{2}{3}X_3.$$

($\bar{X}_1 + \bar{X}_2 + \bar{X}_3 = \pi$ となっていることに注意せよ.)

§4.9 Poisson の少数の法則

$\alpha_1, \alpha_2, \cdots, \alpha_n$ が多数の独立な事象で，α_k のおこる確率 p_k が極めて小さいとし，$\sum_{k=1}^{n} p_k$ がほぼ λ に等しいとする．このとき $\alpha_1, \alpha_2, \cdots, \alpha_n$ のなかでおこるものの数 N の確率法則 P^N は大体パラメータ λ の Poisson 分布に等しいというのが，**Poisson の少数の法則**である．

例えばある工場である特定の日に事故のおこる確率は極めて小さいとし，しかもある日におこったかどうかが，他の日におこるか否かには影響を与えないとしたとき，1年間に事故のおこる回数は大体 Poisson 分布に従うのである．

Poisson の少数の法則を数学的に表現するには，α_k がおこるか否かに応じて $X_k = 1, 0$ とすると，$X_1, X_2, \cdots, X_n$ は独立で，しかも

$$N = X_1 + X_2 + \cdots + X_n$$

となることに注意すればよい．

定理 4.18 (Poisson の少数の法則) おのおのの n に対し

$$X_{n1}, \quad X_{n2}, \quad \cdots, \quad X_{nm} \qquad (m = m(n) \to \infty)$$

は独立で

$$P(X_{nk}=1) = p_{nk}, \qquad P(X_{nk}=0) = 1 - p_{nk}$$

とする．もし

$$\bar{p}_n \equiv \max_k p_{nk} \longrightarrow 0, \qquad p_n \equiv \sum_k p_{nk} \longrightarrow \lambda \qquad (n \to \infty)$$

とすれば

$$N_n = \sum_{k=1}^{m} X_{nk}$$

の確率法則は $n \to \infty$ のとき λ をパラメータとする Poisson 分布に近づく．

証明 $$E(e^{izN_n}) \longrightarrow \exp\{\lambda(e^{iz}-1)\}$$

を示せばよい．

$X_{n1}, X_{n2}, \cdots, X_{nm}$ は独立であるから，

$$E(e^{izN_n}) = \prod_k E(e^{izX_{nk}}) = \prod_k (p_{nk}e^{iz} + (1-p_{nk}))$$
$$= \prod_k (1 + p_{nk}(e^{iz}-1)).$$

中心極限定理の証明と同様に $\alpha_{nk} = p_{nk}(e^{iz}-1)$ とおいて

$$\max_k |\alpha_{nk}| \leqq 2 \max_k p_{nk} \longrightarrow 0,$$

$$\sum_k |\alpha_{nk}| \leqq 2 \sum_k p_{nk} \longrightarrow 2\lambda,$$

$$\sum_k \alpha_{nk} = \sum_k p_{nk}(e^{iz}-1) \longrightarrow \lambda(e^{iz}-1)$$

であるから，補題 4.2 により

$$E(e^{izN_n}) \longrightarrow \exp\{\lambda(e^{iz}-1)\}.$$ ∎

例題 4.9 (i) 定理 4.18 において，$p_{nk} = \lambda/m$ $(m = m(n))$ のとき，定理を初等的に証明せよ．

［ヒント］ N_n の確率法則が二項分布

$$P\{N_n = k\} = \binom{m}{k}\left(\frac{\lambda}{m}\right)^k \left(1-\frac{\lambda}{m}\right)^{m-k}$$
$$= \frac{m(m-1)\cdots(m-k+1)}{m^k}\left(1-\frac{\lambda}{m}\right)^m \left(1-\frac{\lambda}{m}\right)^{-k} \frac{\lambda^k}{k!}$$

であることに注意し，$n \to \infty$ (すなわち $m \to \infty$) のとき，右辺が $e^{-\lambda}\lambda^k/k!$ に近づくことを示せ．

(ii) $X_1, X_2, \cdots, X_k$ が独立で，X_i が Poisson 分布 (パラメータ λ_i) に従うとき，$X_1 + X_2 + \cdots + X_k = n$ のもとにおける $(X_1, X_2, \cdots, X_k)$ の条件付確率法則は多項分布

$$\begin{pmatrix} (n_1, n_2, \cdots, n_k) \\ \dfrac{n!}{n_1!n_2!\cdots n_k!} \displaystyle\prod_{i=1}^{k} \left(\frac{\lambda_i}{\lambda_1+\lambda_2+\cdots+\lambda_n}\right)^{n_i} \end{pmatrix} \qquad \left(\sum_{i=1}^{k} n_i = n\right)$$

に等しいことを示せ．

［ヒント］ $n = \sum n_i$ のときには

$$P\{(X_1, X_2, \cdots, X_k) = (n_1, n_2, \cdots, n_k) \mid \sum_i X_i = n\}$$

$$=\frac{P\{(X_1, X_2, \cdots, X_k)=(n_1, n_2, \cdots, n_k)\}}{P\{\sum_i X_i=n\}}$$

$$=\frac{\prod_i P\{X_i=n_i\}}{P\{\sum_i X_i=n\}}.$$

$\sum_i X_i$ は Poisson 分布(パラメータ $\lambda\equiv\sum\lambda_i$)に従うから

$$=\prod_i e^{-\lambda_i}\frac{\lambda_i^{n_i}}{n_i!}\Big/ e^{-\lambda}\frac{\lambda^n}{n!}$$

$$=\frac{n!}{n_1!n_2!\cdots n_k!}\prod_i\left(\frac{\lambda_i}{\lambda}\right)^{n_i} \qquad (\textstyle\sum n_i=n \text{ に注意}).$$

第5章　確率過程

確率過程は，時とともに変動する偶然現象をあらわすための数学的概念で，確率変数の系

$$\{X_t(\omega), t\in T\} \qquad (\omega\in(\Omega, P))$$

として定義される．ここで T は**時間域**とよばれ，観測時点の集合をあらわしている．普通 T として実数の区間をとるが，実数の離散集合(例えば N)をとることもある．前者は**連続時変数**の確率過程，後者は**離散時変数**の確率過程とよばれる．$X_t(\omega)$ は見本点 ω が実現されたときに，時点 t での観測値であって，ある位相空間 E の中の点であらわされる．E は**状態空間**とよばれる．普通 E としては $\boldsymbol{R}^n$ $(n=1, 2, \cdots)$ をとるが，もっと一般の位相空間でもよい．$\{X_t(\omega), t\in T\}$ を確率過程というときには

"おのおのの t に対して，$X_t(\omega)$ は E 値位相確率変数(すなわち $X_t:\Omega\to E$ が可測 $\mathcal{D}(P)/\mathcal{B}(E)$)"

と仮定する．

これからは，とくに注意しない限り，

$$T \text{ は区間}, \qquad E \text{ は } \boldsymbol{R}^1$$

とする．確率過程を単に**過程**ということも多い．

確率過程は解析学における関数に相当するもので，確率論では最も重要な概念である．解析学に可測関数，連続関数，解析関数というような関数の一般概念があるように，確率論にも，マルチンゲール，加法過程，Markov 過程，定常過程のような確率過程の一般概念がある．また解析学に指数関数，Bessel 関数のような特殊関数があるように，確率論にも Wiener 過程，Poisson 過程のような特殊確率過程がある．さらに微分方程式，超関数に対応して，確率微分方程式，確率超過程がある．しかし解析学における関数の分類や微分方程式の型を直訳しても確率論で興味あるものが得られるわけではない．確率過程の直観的背景である偶然現象の時間的変動を念頭において理論を展開することにより，確率論独特の興

味があり，かつ有用な理論が生れるのである．

§5.1 関数空間 C と D

はじめに準備としてポーランド空間，Kolmogorov σ 加法族について説明する．

ρ を S の上の距離（関数）とする．距離空間 (S,ρ) が完備であるとき，ρ を S の上の**完備距離**といい，(S,ρ) が可分なとき，ρ を**可分距離**という．S の上の完備かつ可分な距離で定まる S の上の位相を**ポーランド位相**という．τ が S の上のポーランド位相のとき，(S,τ) を**ポーランド空間**という．ポーランド空間は位相空間であるが，その位相を定める完備可分距離をそえて完備可分距離空間としてあつかうことができる．可算集合は離散位相に関して，$\boldsymbol{R}^n$ $(n=1,2,\cdots,\infty)$ は普通の位相に関してポーランド空間である．ポーランド空間の可算位相積はポーランド空間である．

位相空間 $S=(S,\tau)$ の Borel 集合族を $\mathcal{B}(S)$ であらわすことは前にのべたが，位相 τ を明示したいときには $\mathcal{B}_\tau(S)$ であらわす．

定理 5.1 S の上にポーランド位相 τ と Hausdorff 位相 τ_1 が与えられていて，τ_1 が τ より弱いならば

$$\mathcal{B}_\tau(S)=\mathcal{B}_{\tau_1}(S). \qquad \text{——}$$

この定理の証明は，解析集合論の知識を要するから，省略する．

集合 T の上の実数値関数の全体を $\boldsymbol{R}^T$ であらわす．$\boldsymbol{R}^T$ の任意の部分集合を T の上の**関数空間**という．T の上の関数空間 F に対し，写像

$$\pi_t:\ F\longrightarrow \boldsymbol{R}^1,\quad f\longmapsto f(t)$$

を t における**射影**という．

F を T の上の関数空間とするとき，$\pi_t^{-1}(\mathcal{B}^1)$ は F の上の σ 加法族である．

$$\mathcal{B}_K(F)=\bigvee_{t\in T}\pi_t^{-1}(\mathcal{B}^1) \qquad (\textstyle\bigvee \text{ については §3.3 を見よ})$$

とおくと，$\mathcal{B}_K(F)$ は，すべての t に対し，π_t が可測 $\mathcal{B}$ となるような σ 加法族 $\mathcal{B}$ の中で最小のものである．$\mathcal{B}_K(F)$ を F の上の **Kolmogorov σ 加法族**という．

以後本節を通じて，T は実数の区間とする．T の上の関数空間の中で確率過程論でとくに重要なのは

$$C=C(T)=T \text{ の上の実数値連続関数の全体},$$

$$D=D(T)=T \text{ の上の第1種不連続関数の全体}$$

である．$f:T\to\boldsymbol{R}^1$ が**第1種不連続**であるとは T の各点（右端を除く）で右連続，かつ T の各点（左端を除く）で有限な左極限が存在することである．

$C=C(T)$ は関数解析で周知の空間である．T がコンパクトなときには，C は距離

$$\rho_u(f,g)=\sup_{t\in T}|f(t)-g(t)|$$

で位相化される．この位相を τ_u であらわす．

$$\rho_u(f_n,f)\longrightarrow 0 \iff f_n \text{ が } f \text{ に一様収束}$$

であるから，τ_u は**一様収束位相**とよばれる．ρ_u は C の上の完備可分な距離であるから，τ_u はポーランド位相である．T がコンパクトでないときには，T は

$$T=\bigcup_n T_n,\qquad T_n \text{ はコンパクト区間},\qquad T_1\subset T_2\subset\cdots$$

の形にかける．

$$\rho_u(f,g)=\sum_{n=1}^\infty 2^{-n}\left[\sup_{t\in T_n}|f(t)-g(t)|\wedge 1\right]$$

とおくと，ρ_u は C の上の完備可分な距離となり，ポーランド位相 τ_u を定める．

$$\rho_u(f_n,f)\longrightarrow 0 \iff f_n \text{ が } f \text{ に広義一様収束}$$

であるから，τ_u は**広義一様収束位相**とよばれる．以後とくに断らない限り，C は τ_u で位相化されたポーランド空間と考える．

$D=D(T)$ は関数解析では考察の対象となったことはないようであるが，確率論では C と同様に重要である．D の上にも，C と同様に ρ_u,τ_u を定義できるが，これはあまり適当でない．まず ρ_u は完備距離であるが，可分ではない．したがって，τ_u は可算開基をもたない．ポーランド位相は可算開基をもつから，τ_u はポーランド位相ではない．ρ_u が可分でないことはつぎの例からわかる．

(＊) $$T=[0,1],\qquad f_\alpha=1_{[0,\alpha)}\quad(\alpha\in T)$$

とすると，$\{f_\alpha,\alpha\in T\}$ は D の中の非可算集合であって，

$$\rho_u(f_\alpha,f_\beta)=1\qquad(\alpha\neq\beta).$$

したがって "α,β が近くても，f_α,f_β は τ_u に関して近くない" から，τ_u は D の上の位相として自然なものとはいえない．以下 T はコンパクトな区間とする．

D に τ_u よりも自然な位相を導入し，しかもその位相に関して，D がポーランド

空間となるようにできれば，理想的であろう．このような位相は A. V. Skorohod によって導入されたので，Skorohod 位相とよばれる．"Skorohod 位相で f が g に近い" というのは，大まかにいうと，"$f(t)$ で t の計り方を少しかえると，$g(t)$ に一様に近くなる" ことである．もっと厳密にいえば，つぎのようになる．T をそれ自身に移す順序保存の位相写像の全体を Φ であらわす．Φ は写像の結合に関して群をなしている．恒等写像 i は Φ の単位元であり，逆写像 φ^{-1} は φ の逆元である．明らかに

$$\Phi \subset C \subset D.$$

また $f, g \in D$ に対し

$$\rho_u(f, g) = \rho_u(f\circ\varphi, g\circ\varphi) \qquad (\varphi \in \Phi)$$

がなりたつ．さて

$$\rho_S(f, g) = \inf_{\varphi\in\Phi}[\rho_u(f\circ\varphi, g)+\rho_u(\varphi, i)]$$

とおくと，ρ_S が D の上の距離となる．ρ_S で定まる位相 τ_S を **Skorohod 位相**という．

$\rho_S(f, g)$ の定義の式で，$\varphi=i$ とおくと

$$\rho_S(f, g) \leqq \rho_u(f, g)$$

が得られるから，τ_S は τ_u よりも弱い位相である．また $\varphi_{\alpha\beta}\in\Phi$ を，$t=0, \beta, 1$ でそれぞれ $0, \alpha, 1$ に等しく，$[0, \beta], [\beta, 1]$ の上でそれぞれ線型となるように定めると，上の例(＊)の $\{f_\alpha=1_{[0,\alpha]}\}$ に関して

$$f_\alpha\circ\varphi_{\alpha\beta}=f_\beta, \ \rho_S(f_\alpha, f_\beta) \leqq \rho_u(\varphi_{\alpha\beta}, i) \leqq |\alpha-\beta|$$

が得られるから，α, β が近いときには，f_α, f_β は τ_S に関して近くなる．

ρ_S は可分であるが，完備ではない(例題5.1参照)．しかし τ_S はポーランド位相である．これを示すには，ρ_S と位相的には同等な完備可分な距離を定義すればよい．下にのべる **Billingsley の距離** ρ_B がこの性質をもっている．要点をわかりやすくするため，T がコンパクトなときに説明しよう．$\varphi\in\Phi$ に対し

$$\lambda(\varphi) = \sup_{t\neq s}\left|\log\frac{\varphi(t)-\varphi(s)}{t-s}\right|$$

とおく．$\lambda(\varphi)$ が小さいというのは，$\varphi(t)$ の傾きが一様に1に近いことを意味する．これは $\varphi(t)$ が '一様に t に近い' というよりも強い条件である．

$$\Psi = \{\varphi\in\Phi \mid \lambda(\varphi)<\infty\}$$

は Φ の部分群で

$$\lambda(\varphi^{-1}) = \lambda(\varphi), \qquad \lambda(\varphi\circ\psi) \leqq \lambda(\varphi)+\lambda(\psi).$$

さて

$$\rho_B(f, g) = \inf_{\psi\in\Psi}[\rho_u(f\circ\psi, g)+\lambda(\psi)]$$

とおくと，ρ_B は D の上の完備可分な距離となり，しかも ρ_S と同じ位相 τ_S を与える．この証明は技術的に複雑であるから，省略する．

以後 $D=D(T)$ は τ_S で位相化し，ポーランド空間と見なす．

定理 5.2　(i)　$\mathcal{B}(C) = \mathcal{B}_K(C)$,

(ii)　$\mathcal{B}(D) = \mathcal{B}_K(D)$.

証明　T をコンパクトな区間と仮定して証明するが，一般のばあいの証明も本質的に同じである．

(i)　射影 $\pi_t: C\to \boldsymbol{R}^1$ は連続であるから，もちろん可測 $\mathcal{B}(C)$ である．したがって

$$\mathcal{B}_K(C) \subset \mathcal{B}(C).$$

逆の包含関係をいう．$g\in C$ を任意に固定して

$$F_g(f) = \rho_u(f, g) \equiv \sup_{t\in T}|f(t)-g(t)|$$

とおくと，f, g の連続性により，

$$F_g(f) = \sup_{t\in T\cap \boldsymbol{Q}}|f(t)-g(t)| = \sup_{t\in T\cap \boldsymbol{Q}}|\pi_t(f)-\pi_t(g)|.$$

π_t が可測 $\mathcal{B}_K(C)$ であるから，$F_g: C\to\boldsymbol{R}^1$ も同様．したがって

$$U(g, r) = \{f\in C \mid \rho_u(f, g)<r\} = \{f\in C \mid F_g(f)<r\} \in \mathcal{B}_K(C).$$

ρ_u が可分であるから，$\mathcal{B}(C)$ は $U(g, r)$, $g\in C$, $r>0$ で生成される．これから $\mathcal{B}(C)\subset\mathcal{B}_K(C)$ が得られ，結局 $\mathcal{B}(C)$ と $\mathcal{B}_K(C)$ とが一致することがわかる．

(ii)　T がコンパクトであるから，$D=D(T)$ の元は有界かつ Borel 可測(したがって可積分)な関数である．

$$\rho_\lambda(f, g) = \int_T|f(t)-g(t)|dt$$

は D の上の Hausdorff 位相を定める．この位相を τ_λ であらわす．位相 τ_λ に関する D の上の Borel 集合族を $\mathcal{B}_\lambda(D)$ であらわす．さて

$$\mathcal{B}(D)=\mathcal{B}_\lambda(D),\qquad \mathcal{B}_\lambda(D)=\mathcal{B}_K(D)$$

をいえば，$\mathcal{B}(D)=\mathcal{B}_K(D)$ となり証明はおわる．

まず $\mathcal{B}(D)=\mathcal{B}_\lambda(D)$ を示そう．τ_S がボーランド位相であるから，定理5.1により，τ_λ が τ_S より弱いことをいえば，すなわち，

$$\rho_S(f_n,f)\longrightarrow 0 \Longrightarrow \rho_\lambda(f_n,f)\longrightarrow 0$$

をいえば十分である．仮定から，$\varphi_n\in\Phi$ が存在して

$$\rho_u(f_n,f\circ\varphi_n)\longrightarrow 0,\qquad \rho_u(\varphi_n,i)\longrightarrow 0$$

となる．ゆえに t に関して一様に

$$f_n(t)-f(\varphi_n(t))\longrightarrow 0,\qquad \varphi_n(t)\longrightarrow t.$$

$f\ (\in D)$ は有界関数であるから，上の第1式から $\{f_n\}$ が一様に有界な関数列となる．また第2式から，f の連続点で

$$f(\varphi_n(t))-f(t)\longrightarrow 0.$$

これと上の第1式とを組み合せて，f の連続点 t で

$$f_n(t)\longrightarrow f(t)$$

がでる．$f\ (\in D)$ の不連続点は可算個であるから，有界収束定理により

$$\rho_\lambda(f_n,f)=\int_T|f_n(t)-f(t)|dt\longrightarrow 0.$$

これで $\mathcal{B}(D)=\mathcal{B}_\lambda(D)$ が証明された．

残るのは $\mathcal{B}_\lambda(D)=\mathcal{B}_K(D)$ の証明である．$[a,b]\subset T$ に対し

$$I_{a,b}(f)=\int_a^b f(t)dt$$

とおくと，

$$|I_{a,b}(f)-I_{a,b}(g)|\leqq\int_a^b|f(t)-g(t)|dt\leqq\rho_\lambda(f,g)$$

により，$I_{a,b}:D\to\boldsymbol{R}^1$ は τ_λ に関し連続である．$f\ (\in D)$ の右連続性により

$$\pi_t(f)=f(t)=\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}I_{a,a+\varepsilon}(f).$$

これは $\pi_t:D\to\boldsymbol{R}^1$ は可測 $\mathcal{B}_\lambda(D)$ であることを意味するから，

$$\mathcal{B}_K(D)\subset\mathcal{B}_\lambda(D)$$

が得られた．ρ_λ は D の上の可分距離であるから，$\mathcal{B}(C)\subset\mathcal{B}_K(C)$ の証明と同様に考えて，任意の $g\in D$ に対し，

$$F(f)=\rho_\lambda(f,g)=\int_T|f(t)-g(t)|dt$$

が可測 $\mathcal{B}_K(D)$ となることをいえば，

$$\mathcal{B}_\lambda(D)\subset\mathcal{B}_K(D)$$

が得られ，証明が完了する．

$$F_n(f)=\sum_{i=1}^n|f(t_{n,i})-g(t_{n,i})|(t_{n,i}-t_{n,i-1})$$

$$(\{t_{n,0},t_{n,1},\cdots,t_{n,n}\}\text{ は } T \text{ の } n \text{ 等分点})$$

とおくと，$f(t_{n,i})=\pi_{t_{n,i}}(f)$ により，$F_n:D\to\boldsymbol{R}^1$ は可測 $\mathcal{B}_K(D)$ である．$f,g\ (\in D)$ は Riemann 可積分であるから，

$$F(f)=\lim_{n\to\infty}F_n(f).$$

これは F が可測 $\mathcal{B}_K(D)$ であることを意味する．∎

例題 5.1　$D=D(T)$ $(T=[0,1])$ について下の命題を証明せよ．

(i) $f\in D$ は右連続階段関数列の一様収束極限である．また逆も真である．

(ii) $f_n=1_{[0,1/n)}$ $(n=1,2,\cdots)$ は D の中の ρ_S に関する Cauchy 列であるが，ρ_S に関して収束しない．また f_n は ρ_B に関しては Cauchy 列ではない．

(iii) 上述の $\{f_n\}$ は ρ_B に関しては Cauchy 列ではない．

(iv) $\rho_B,\rho_S,\rho_\lambda$ は $(D$ の中の$)$ 距離の条件をみたす．

(v) C の上の τ_u は D の上の τ_S の C への相対化である．

［ヒント］　$f_1,f_2,\cdots,f\in C$ のとき $\rho_S(f_n,f)\to 0$ と $\rho_u(f_n,f)\to 0$ とが同等なことをいえばよい．

§5.2　確率過程に関する一般事項

$\{X_t,t\in T\}$ を確率過程とする．本章の冒頭に注意したように，T は実数の区間，X_t は実確率変数である．実確率変数の全体 $L^0=L^0(\Omega,P)$ は距離

$$\rho_0(X,Y)=E(|X-Y|\wedge 1)$$

に関して完備可分距離空間である．また T も通常の距離

$$d(t,s)=|t-s|$$

に関して可分距離空間になっている．したがって確率過程 $\{X_t,t\in T\}$ に対して，距離空間 T から完備可分距離空間 $L^0=(L^0,\rho_0)$ の中への写像 $t\mapsto X_t$ が考えられ

る．この写像が連続なときには，$\{X_t, t\in T\}$ は**確率連続**であるといい，この写像が一様連続なときには，$\{X_t, t\in T\}$ は**一様確率連続**という．明らかに一様確率連続ならば，確率連続であるが，T がコンパクトであれば，逆もまた真である．

§3.7で示した不等式：

$$\varepsilon P\{|X-Y|>\varepsilon\} \leqq \rho_0(X, Y) \leqq \varepsilon+P\{|X-Y|>\varepsilon\} \qquad (0<\varepsilon<1)$$

を用いると，$\{X_t, t\in T\}$ が確率連続であるとは，

$$(\forall s\in T)(\forall\varepsilon>0)\quad \lim_{t\to s} P\{|X_t-X_s|>\varepsilon\}=0$$

がなりたつことと定義してもよいことがわかる．また，一様確率連続は

$$(\forall\varepsilon>0)\quad \lim_{\delta\to 0}\sup_{\substack{|t-s|<\delta\\ t,s\in T}} P\{|X_t-X_s|>\varepsilon\}=0$$

と定義してもよい．

上述の考察では，確率過程 $\{X_t, t\in T\}$ に対して，写像 $t\mapsto X_t$ を考えたが，つぎに別の観点から考察してみよう．$X_t=X_t(\omega)$ は t と ω との関数であるから，ω を固定すると，t の関数が得られる．この関数を $X.(\omega)$ または $(X_t(\omega), t\in T)$ であらわし，ω に対する $\{X_t, t\in T\}$ の**見本過程**（または**見本関数**）という．これは，直観的には，見本 ω が実現されたとき，問題の偶然現象がいかなる時間的変動をするかをあらわしている．すべての ω に対して，$X.(\omega)$ は $\boldsymbol{R}^T$ の元である．とくにすべての ω に対して，$X.(\omega)$ が $C=C(T)$ に属するときには，$\{X_t(\omega), t\in T\}$ を（**見本**）**連続過程**，または簡単に **C 過程**という．同様にすべての ω に対して $X.(\omega)$ が $D=D(T)$ に属するときには，（**見本**）**第1種不連続過程**または簡単に **D 過程**という．

定理 5.3　(i)　$\{X_t, t\in T\}$ が C 過程のときには，見本過程 $X.(\omega)$ は，ω の C 値関数と見て，C 値位相確率変数である．

(ii)　$\{X_t, t\in T\}$ が D 過程のときには，見本過程 $X.(\omega)$ は，ω の D 値関数と見て，D 値位相確率変数である．

証明　(i), (ii) は同じ方法で証明されるから，(i) のみ証明する．C はボーランド空間であるから，可算開基をもつ．したがって $X.(\omega)$ が C 値位相確率変数であることを示すには

$$X.^{-1}(\mathcal{B}(C))\subset\mathcal{D}(P)$$

をいえばよい．個々の $t\in T$ に対しては，

$$X_t^{-1}(\mathcal{B}^1)\subset\mathcal{D}(P)$$

となることは明らかである．$\pi_t: C\to\boldsymbol{R}^1$ を射影写像とすると，$X_t=\pi_t\circ X.$ であるから，上の式は

$$X.^{-1}(\pi_t^{-1}(\mathcal{B}^1))\subset\mathcal{D}(P), \qquad t\in T$$

を意味する．さて

$$\mathcal{B}=\{B\subset C\,|\,X.^{-1}(B)\in\mathcal{D}(P)\}$$

とすれば，$\mathcal{B}$ は σ 加法族であり，上に証明した式から

$$\mathcal{B}\supset\pi_t^{-1}(\mathcal{B}^1), \qquad t\in T,$$

したがって

$$\mathcal{B}\supset\mathcal{B}_K(C)=\mathcal{B}(C) \qquad (\text{定理 } 5.2\text{ (i)}).$$

これから

$$X.^{-1}(\mathcal{B}(C))\subset X.^{-1}(\mathcal{B})\subset\mathcal{D}(P). \qquad \blacksquare$$

つぎの定理は上の定理の逆もまた真であることを示している．

定理 5.4　(i)　$X(\omega)$ が C 値位相確率変数 $(C=C(T))$ であるとき，

$$X_t(\omega)=\pi_t(X(\omega)), \qquad t\in T \qquad (\pi_t: C\to\boldsymbol{R}\text{ は射影})$$

は C 過程で，その見本過程 $X.(\omega)$ は $X(\omega)$ と一致する．

(ii)　上の命題で C を D でおきかえてもなりたつ．

証明　(i) だけを証明する．個々の $t\in T$ に対して，$X_t(\omega)$ が P 可測となることをいえば，あとはほとんど明らかである．任意の t に対し

$$X_t^{-1}(\mathcal{B}^1)=X^{-1}(\pi_t^{-1}(\mathcal{B}^1))\subset X^{-1}(\mathcal{B}_K(C))=X^{-1}(\mathcal{B}(C))$$
$$(\text{定理 } 5.2\text{ (i)}).$$

$X(\omega)$ は C 値位相確率変数であるから，

$$X^{-1}(\mathcal{B}(C))\subset\mathcal{D}(P), \qquad \text{したがって}\quad X_t^{-1}(\mathcal{B}^1)\subset\mathcal{D}(P).$$

これは $X_t(\omega)$ が P 可測となることを意味する．$\blacksquare$

上の2定理により，C 過程と C 値位相確率変数とは見方が違うだけで，本質的には同じものであるということができる．D 過程と D 値位相確率変数についても同様である．

確率過程 $\{X_t, t\in T\}$ の第3の観点は，$X_t(\omega)$ を2変数 (t, ω) の実関数とみなすことである．この関数（定義域は $T\times\Omega$）が可測 $\mathcal{B}(T)\times\mathcal{D}(P)$ のときには，与えられた確率過程を**可測過程**とよぶ．

定理 5.5 (i) 2変数 $(t,\omega)\in T\times\Omega$ の実関数 $X_t(\omega)$ が可測 $\mathcal{B}(T)\times\mathcal{D}(P)$ ならば，$\{X_t(\omega), t\in T\}$ は可測過程である．

(ii) $\{X_t, t\in T\}$ が可測過程ならば，すべての ω に対して $X_t(\omega)$ は t の関数と見て可測 $\mathcal{B}(T)$ である．

(iii) C 過程も D 過程も可測過程である．

証明 (i) で証明すべき唯一の点は“すべての t に対して，$X_t(\omega)$ が ω の関数と見て可測 $\mathcal{D}(P)$ である”ということであるが，これは例題 5.2 (i) から明らかである．(ii) も同様．(iii) を，D 過程について $T=[0,1]$ のときに，証明する．$\{X_t, t\in T\}$ を D 過程とせよ．

$$X_t^{(n)}(\omega)=\sum_{k=1}^{n-1}1_{[(k-1)/n,k/n)}(t)X_{k/n}(\omega)+1_{\{1\}}(t)X_1(\omega)$$

とおくと，すべての ω に対して，$X_t(\omega)$ は t の右連続関数であるから，すべての (t,ω) に対し

$$X_t(\omega)=\lim_{n\to\infty}X_t^{(n)}(\omega).$$

上の和式の各項は明らかに (t,ω) の関数と見て可測 $\mathcal{B}(T)\times\mathcal{D}(P)$ であるから，$X_t^{(n)}(\omega)$ も同様．したがって $X_t(\omega)$ も同様である．∎

$\{X_t, t\in T\}$, $\{Y_t, t\in T\}$ を確率過程とする．

$$\text{“}(\forall t)\quad P\{X_t=Y_t\}=1\text{”}$$

のとき，$\{X_t\}$, $\{Y_t\}$ は**同等**であるという．また

$$\text{“}P\{(\forall t)\ \ X_t=Y_t\}=1\text{”}$$

のとき，$\{X_t\}$, $\{Y_t\}$ は**強同等**であるという．最後に

“任意の有限個の時点 $t_1, t_2, \cdots, t_n\in T$ に対し $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ と $(Y_{t_1}, Y_{t_2}, \cdots, Y_{t_n})$ とが同じ確率法則をもつ”

とき，$\{X_t\}$, $\{Y_t\}$ は**法則同等**であるという．明らかに

$$\text{強同等}\implies\text{同等}\implies\text{法則同等}.$$

注意 C 過程の定義で“すべての ω に対して，$X.(\omega)$ が $C=C(T)$ に属する”といったが，もう少しゆるくして“**ほとんど**すべての ω に対して，$X.(\omega)\in C$”といってもよい．このときには，(t に無関係な) ある P 零集合の上で，すべての t に対し $X_t(\omega)$ を修正してもとのばあいに帰着できる．このようなゆるい意味での連続過程では $X.(\Omega)$ は正確には C に属するとはいえないが，$X.(\omega)$ の確率法則 $P^{X.}$ は C に集中している (すなわち $P^{X.}(C)$ $=1$) から，$P^{X.}$ を C の上の確率測度とみなすことができる．このように定義をゆるめておけば，C 過程と強同等な確率過程は C 過程である．

例題 5.2 (i) $\mathcal{S}$ を S の上の σ 加法族，$\mathcal{T}$ を T の上の σ 加法族とする．f: $S\times T\to \boldsymbol{R}^1$ が可測 $\mathcal{S}\times\mathcal{T}$ ならば，任意の $s\in S$ に対して

$$f^{(s)}:\ T\longrightarrow \boldsymbol{R}^1,\quad t\longmapsto f(s,t)$$

は可測 $\mathcal{T}$ であること，任意の $t\in T$ に対して

$$f^{(t)}:\ S\longrightarrow \boldsymbol{R}^1,\quad s\longmapsto f(s,t)$$

は可測 $\mathcal{S}$ であることを示せ．

[ヒント] Fubini の定理の証明にこの事実は含まれている．“$f=1_{A\times B}$ $(A\in\mathcal{S}, B\in\mathcal{T})$ に対しては結論がなりたつ．”また“$f=1_C$ $(C\subset S\times T)$ に対して結論がなりたつような C の全体 $\mathcal{C}$ は Dynkin 族をなす．”このことから，“$f=1_C$ $(C\in\mathcal{S}\times\mathcal{T})$ に対しては結論がなりたつ”ことがわかる．任意の f (可測 $\mathcal{S}\times\mathcal{T}$) はこのような 1_C の1次結合の列の極限であることに注意すれば，証明は完了する．

(ii) C 過程 $\{X_t, t\in T\}$, $\{Y_t, t\in T\}$ が法則同等ならば，両者の見本過程の確率法則 $P^{X.}$, $P^{Y.}$ は一致することを示せ．D 過程についても同様．

[ヒント] C 過程についていう．$X.$, $Y.$ は C 値位相確率変数で，C はポーランド空間 (したがって適当な距離で完備可分距離空間) であるから，$P^{X.}$, $P^{Y.}$ は正則である．したがって両者が $\mathcal{B}(C)$ の上で一致することをいえばよい．法則同等の仮定により

$$A=\pi_{t_1}^{-1}(B_1)\cap\pi_{t_2}^{-1}(B_2)\cap\cdots\cap\pi_{t_n}^{-1}(B_n)\qquad(B_i\in\mathcal{B}^1,\ i=1,2,\cdots,n)$$

に対しては $P^{X.}(A)=P^{Y.}(A)$．Dynkin 族定理を用いて，$A\in\mathcal{B}(C)$ $(=\mathcal{B}_K(C)$, 定理 5.2 (i)) に対しても，この等式がなりたつ．

§5.3 情報と増大情報系

はじめに閉 σ 加法族という概念を導入しよう．(Ω, P) を基礎の確率空間とすると，P 可測集合の全体 $\mathcal{D}(P)$ は σ 加法族である．$\mathcal{D}(P)$ の部分 σ 加法族 ($\mathcal{D}(P)$ の部分族で Ω の上の σ 加法族となっているもの) をここでは簡単に σ 加法族とよぶことにする．すべての P 零集合を含む σ 加法族を**閉 σ 加法族**といい，閉 σ 加法族の全体を $\Phi=\Phi(\Omega,P)$ であらわす．Φ は σ 加法族に関する演算 $\vee$, $\wedge$ に関して閉じている．σ 加法族 $\mathcal{B}$ に対し，$\mathcal{B}\vee\boldsymbol{2}$ は $\mathcal{B}$ を含む最小の閉 σ 加法族で

ある.

X を S 値確率変数 (S は抽象空間) とするとき, $\mathcal{F}$ ($\in \Phi$) に関して X が可測である (X が $\mathcal{F}$ **可測**または**可測** $\mathcal{F}$ ともいう) とは

$$X^{-1}(\mathcal{D}(P^X)) \subset \mathcal{F} \quad \text{すなわち} \quad X \in \mathcal{F}/\mathcal{D}(P^X)$$

となることである. とくに X が S 値位相確率変数 (S は位相空間) のときには, これは §3.2 にのべた意味の '可測 $\mathcal{F}$' すなわち

$$X^{-1}(\mathcal{B}(S)) \subset \mathcal{F} \qquad (\text{すなわち } X \in \mathcal{F}/\mathcal{B}(S))$$

と同等であることは,

$$X^{-1}(\mathcal{B}(S)) \subset X^{-1}(\mathcal{D}(P^X)) \subset X^{-1}(\mathcal{B}(S)) \vee \boldsymbol{2}$$

から明らかである.

さて

$$\mathcal{F}[X] = X^{-1}(\mathcal{D}(P^X)) \vee \boldsymbol{2}$$

とおくと, $\mathcal{F}[X]$ は閉 σ 加法族であり, X は可測 $\mathcal{F}[X]$ である. しかもこのような閉 σ 加法族の中で最小のものである. $\mathcal{F}[X]$ を X で生成される**閉 σ 加法族**という. 確率変数系 $\{X_\lambda, \lambda \in \Lambda\}$ については

$$\mathcal{F}[X_\lambda, \lambda \in \Lambda] = \bigvee_\lambda \mathcal{F}[X_\lambda]$$

をこの系で生成される**閉 σ 加法族**という. これはすべての X_λ, $\lambda \in \Lambda$ を可測 $\mathcal{F}$ とする閉 σ 加法族 $\mathcal{F}$ の中で最小のものである.

定理 5.6 任意の閉 σ 加法族 $\mathcal{F}$ はある確率変数 X で生成される. 実は X として実確率変数をとることができる.

証明 本講では P を可分完全確率測度と仮定したことを思いおこすと, $L^1 = L^1(\Omega, P)$ は可分距離空間である. したがって, その部分空間

$$L_{\mathcal{F}}^1 = \{X \in L^1 \mid X \text{ は可測 } \mathcal{F}\}$$

も可分距離空間である. この中の可算稠密集合を $\{X_1, X_2, \cdots\}$ とすると, X_i は可測 $\mathcal{F}$ であるから,

$$\mathcal{F}[X_1, X_2, \cdots] \subset \mathcal{F}.$$

また任意の $A \in \mathcal{F}$ に対し, $1_A \in L_{\mathcal{F}}^1$ であるから, $\{X_n\}$ の適当な部分列 $\{Y_n\}$ をとって

$$E(|Y_n - 1_A|) \longrightarrow 0 \qquad (n \to \infty).$$

$\{Y_n\}$ の部分列 $\{Z_n\}$ を適当にとれば,

$$Z_n \longrightarrow 1_A \text{ a.s.}$$

これは $\mathcal{F} \subset \mathcal{F}[X_1, X_2, \cdots]$ を示し, 結局

$$\mathcal{F}[X_1, X_2, \cdots] = \mathcal{F}$$

となる. これは $\mathcal{F}$ が $\boldsymbol{R}^\infty$ 値位相確率変数 $(X_1, X_2, \cdots)$ で生成される閉 σ 加法族であることを意味する. $\boldsymbol{R}^\infty$ には可算分離族 $\{E_1, E_2, \cdots\}$ が存在するから, E_n の指示関数を e_n とし,

$$X = \sum_{n=1}^{\infty} 3^{-n} e_n(X_1, X_2, \cdots)$$

とおくと, X は実確率変数で,

$$\mathcal{F}[X] = \mathcal{F}[X_1, X_2, \cdots] = \mathcal{F}$$

となることが容易にわかる. ▮

上の定理により, 閉 σ 加法族 $\mathcal{F}$ は $\mathcal{F}[X]$ の形にかけるわけであるが, このような X はいくらもある. 確率変数 X, Y に対して

$$Y(\omega) = \varphi(X(\omega)) \text{ a.s.}$$

となる写像 $\varphi: X(\Omega) \to Y(\Omega)$ が存在するとき,

$$Y \prec_i X \text{ a.s.}$$

であらわす. これは直観的には Y の与える情報 (information) が X のそれよりも少ないことを意味し, 添字の i は information の頭文字である. '$Y \prec_i X$ a.s.' かつ '$X \prec_i Y$ a.s.' のとき

$$Y \sim_i X \text{ a.s.}$$

であらわす.

定理 5.7 (i) $Y \prec_i X$ a.s. $\Longleftrightarrow$ $\mathcal{F}[Y] \subset \mathcal{F}[X]$,

(ii) $Y \sim_i X$ a.s. $\Longleftrightarrow$ $\mathcal{F}[Y] = \mathcal{F}[X]$.

証明 (ii)は(i)からすぐにでるから, (i)のみを証明しよう. $Y \prec_i X$ a.s. とせよ. 定義により写像 $\varphi: X(\Omega) \to Y(\Omega)$ が存在して

$$Y(\omega) = \varphi(X(\omega)) \text{ a.s.}$$

となる. φ を拡張して, $\varphi: S \to T$ (S, T はそれぞれ X, Y の値域空間) とすることができる. 上の等式の例外の P 零集合を N とし

$$\Omega_1 = \Omega - N, \quad P_1 = P|_{\Omega_1}, \quad X_1 = X|_{\Omega_1}, \quad Y_1 = Y|_{\Omega_1}$$

とし，Ω_1 の上で考えると，例外点なしに

$$Y_1(\omega)=\varphi(X_1(\omega))$$

となるから，

$$\begin{aligned}&A=Y_1^{-1}(E),\quad E\in\mathcal{D}(P^{Y_1})\\ &\Longrightarrow A=X_1^{-1}(\varphi^{-1}(E))\\ &\Longrightarrow F\equiv\varphi^{-1}(E)\in\mathcal{D}(P^{X_1}),\quad A=X_1^{-1}(F).\end{aligned}$$

これから

$$Y_1^{-1}(\mathcal{D}(P^{Y_1}))\subset X_1^{-1}(\mathcal{D}(P^{X_1}))$$

が得られる．したがって，$\omega\in\Omega-N$ では $X(\omega)=X_1(\omega)$，$Y(\omega)=Y_1(\omega)$ であるから，

$$Y^{-1}(\mathcal{D}(P^Y))\vee\boldsymbol{2}\subset X^{-1}(\mathcal{D}(P^X))\vee\boldsymbol{2}\quad すなわち\quad \mathcal{F}[Y]\subset\mathcal{F}[X].$$

逆に $\mathcal{F}[Y]\subset\mathcal{F}[X]$ とせよ．T の上の可算分離族を $\{E_n\}$ とすると，

$$Y^{-1}(E_n)\in\mathcal{F}[Y]\subset\mathcal{F}[X]=X^{-1}(\mathcal{D}(P^X))\vee\boldsymbol{2},$$

したがって

$$Y^{-1}(E_n)=X^{-1}(F_n)\cup N_n,\qquad F_n\in\mathcal{D}(P^X),\quad P(N_n)=0$$

とかける．$N=\bigcup_n N_n$ $(P(N)=0)$ とおき，上と同様に Ω_1, P_1, X_1, Y_1 を定義し，Ω_1 の上で考えると，上の式から

$$Y_1^{-1}(E_n)=X_1^{-1}(F_n),\qquad n=1,2,\cdots$$

となる．任意の $t\in T$ は

$$\{t\}=\bigcap_n E_n^{\varepsilon_n}\qquad(\varepsilon_n=1,0;\ E^1=E,\ E^0=E^c)$$

とかけるから，

$$Y_1^{-1}(t)=X_1^{-1}(\varDelta_t),\qquad \varDelta_t=\bigcap_n F_n^{\varepsilon_n}\qquad(\varepsilon_n=1,0;\ F^1=F,\ F^0=F^c).$$

明らかに

$$t\neq t'\Longrightarrow \varDelta_t\cap\varDelta_{t'}=\phi,\qquad \bigcup_{t\in T_1}\varDelta_t=S_1$$

であるから，写像 $\varphi_1: S_1\to T_1$ を

$$\varphi_1(s)=t,\qquad s\in\varDelta_t$$

と定義すると，

$$Y_1(\omega)=\varphi_1(X_1(\omega)),\qquad \omega\in\Omega_1$$

となる．φ_1 を任意に拡張して $\varphi: S\to T$ を定めると，

$$Y(\omega)=\varphi(X(\omega)),\qquad \omega\in\Omega-N$$

すなわち，$Y(\omega)=\varphi(X(\omega))$ a.s. となる．▌

さて条件 '$Y\prec_i X$ a.s.' は次の条件と同等である．

"P 零集合 N が存在して

$\omega_1,\omega_2\in\Omega-N$　のときには　$X(\omega_1)=X(\omega_2)\Longrightarrow Y(\omega_1)=Y(\omega_2)$."

したがって '$Y\prec_i X$ a.s.' は "$X(\omega)$ の値がわかれば，ほとんど確実に $Y(\omega)$ の値がきまる" ことを意味する．この事実を念頭において，'$Y\prec_i X$ a.s.' のとき，Y の与える情報 (information) は X の与える情報よりも少ない (正確には '多くない' というべきであろう) という．したがって，'$Y\sim_i X$ a.s.' のときには，Y の与える情報は X の与える情報に等しいという．定理 5.7 (ii) により '$Y\sim_i X$ a.s.' は '$\mathcal{F}[Y]=\mathcal{F}[X]$' と同等であるから，$\mathcal{F}[X]$ を 'X の与える**情報**' と定義してよいであろう．この定義によると，定理 5.7 (i) により $\mathcal{F}[Y]\subset\mathcal{F}[X]$ は Y の与える情報は X の与える情報よりも少ないことを意味する．定理 5.6 によれば，任意の閉 σ 加法族 $\mathcal{F}$ は必ず $\mathcal{F}[X]$ の形にかけるから，$\mathcal{F}$ を**情報**とよぶことにしてもよいであろう．以上の考察を背景として次の定義を与える．

任意の閉 σ 加法族 $\mathcal{F}$ を**情報**とよび，$\mathcal{F}_1\subset\mathcal{F}_2$ のとき情報 $\mathcal{F}_1$ は情報 $\mathcal{F}_2$ よりも小さい (正確には '大きくない') という．確率変数 X によって生成される閉 σ 加法族を X の与える情報といい，$\mathcal{F}[X]$ であらわす．同様に確率変数系 X_λ, $\lambda\in\Lambda$ の与える情報 $\mathcal{F}[X_\lambda,\lambda\in\Lambda]$ を定義する．また条件付確率 $P(A|\mathcal{F})$ を情報 $\mathcal{F}$ のもとにおける (条件付) 確率といい，条件付平均値 $E(X|\mathcal{F})$ を情報 $\mathcal{F}$ のもとにおける (条件付) 平均値という．$\mathcal{F}[Y]\subset\mathcal{F}$ のとき，すなわち Y が可測 $\mathcal{F}$ のときには，Y の値が情報 $\mathcal{F}$ で定まるという．

情報が時の経過につれて増大していく状況をあらわす概念として，増大情報系を導入する．T を実数の区間または離散集合とし，時間域とよぶことは確率過程のときと同様である．時点 $t\in T$ に依存する情報の系 $\boldsymbol{F}=\{\mathcal{F}_t, t\in T\}$ が

増大性　　$s<t\Longrightarrow\mathcal{F}_s\subset\mathcal{F}_t$

をみたすとき，$\boldsymbol{F}$ を**増大情報系**という．ただし，T が区間のときには，さらに

右連続性　　$\mathcal{F}_t=\mathcal{F}_{t+}\qquad\left(\mathcal{F}_{t+}=\bigwedge_{s>t}\mathcal{F}_s\right)$

の仮定も要求することにする．

$\{\mathcal{F}_t, t\in T\}$ (T は区間) が増大性のみをみたすときには,

$$\{\mathcal{F}_{t+}, t\in T\}$$

は増大性も右連続性もみたすから，上記の意味で増大情報系となる．これをもとの情報系の**右連続化**という．

確率過程 $X=\{X_t, t\in T\}$ に対して

$$\mathcal{F}_t=\begin{cases}\mathcal{F}[X_s, s\leqq t] & (T\text{ が離散のとき}),\\ \bigcap_{s>t}\mathcal{F}[X_u, u\leqq s] & (T\text{ が区間のとき})\end{cases}$$

とおくと，増大情報系 $\{\mathcal{F}_t, t\in T\}$ が得られる．これを X によって**生成**される情報系といい $\boldsymbol{F}(X)=\{\mathcal{F}_t[X], t\in T\}$ であらわす．

$\boldsymbol{F}=\{\mathcal{F}_t, t\in T\}$ を情報系，$X=\{X_t, t\in T\}$ を確率過程とする．すべての $t\in T$ に対し，X_t が可測 $\mathcal{F}_t$ のとき，X は $\boldsymbol{F}$ に**適合**するという．この条件は

$$\mathcal{F}_t[X]\subset\mathcal{F}_t,\quad t\in T$$

と同等である．

例題 5.3　(i)　$\{X_t, t\in T\}$ が確率連続ならば，T の任意の稠密部分集合 T' に対して

$$\mathcal{F}[X_t, t\in T]=\mathcal{F}[X_t, t\in T']$$

となることを示せ．

［ヒント］　右辺を $\mathcal{F}'$ とする．任意の $t\in T$ に対して，$X_t=\underset{n\to\infty}{\text{l.i.p.}}\,X_{t_n}$，$t_n\in T'$．部分列をとって $X_t=\lim_{n\to\infty}X_{s_n}$ a.s.，$s_n\in T'$．X_{s_n} は可測 $\mathcal{F}'$ であるから，X_t も同様．

(ii)　$\{X_t, t\in T\}$ が C 過程(または D 過程)ならば

$$\mathcal{F}[X_t, t\in T]=\mathcal{F}[X.]$$

となることを示せ．

［ヒント］　$\mathcal{B}_K(C)=\mathcal{B}(C)$，$\mathcal{B}_K(D)=\mathcal{B}(D)$ を用いよ．

§5.4　停止時

保険を解約したり，計画を中止したり，賭博をやめたりするばあいには，時点 n でやめるかどうかは，その時点までの情報 $\mathcal{F}_n$ によってきめなければならない．持金がなくなったときに賭けをやめることはできるが，最後に大損をして“最後の賭けはしなかったことにしてほしい”といっても，ききいれられない．この事

情を数学的にいうと，つぎのようになる．

$\boldsymbol{F}=\{\mathcal{F}_1, \mathcal{F}_2, \cdots\}$ を増大情報系(離散時変数)とする．$\{1, 2, 3, \cdots, \infty\}$ の中の値をとる確率変数 $\tau=\tau(\omega)$ が $\boldsymbol{F}$ に関する**停止時**(やめどき，stopping time)であるとは，

$$\{\tau=n\}\in\mathcal{F}_n,\quad n=1, 2, \cdots$$

がなりたつことである．$\{\tau=n\}$ $(n=1, 2, \cdots)$ は時点 n でやめることを意味し，$\{\tau=\infty\}$ はいつまでもやめないことを意味する．

τ が $\boldsymbol{F}$ に関する停止時であれば，$\tau+m$ $(m=1, 2, \cdots)$ も同様である．なぜならば，

$$n\leqq m\quad\text{のときには}\quad\{\tau+m=n\}=\phi\in\mathcal{F}_n,$$
$$n>m\quad\text{のときには}\quad\{\tau+m=n\}=\{\tau=n-m\}\in\mathcal{F}_{n-m}\subset\mathcal{F}_n.$$

しかし，$\tau-m$ $(m=1, 2, \cdots)$ は一般に $\boldsymbol{F}$ に関する停止時ではない．

$\tau(\omega)\equiv k$ (定数) も $\boldsymbol{F}$ に関する停止時である．なぜならば

$$n=k\text{ または }n\neq k\quad\text{に応じて}\quad\{\tau=n\}=\Omega\text{ または }\phi\in\mathcal{F}_n.$$

τ が $\boldsymbol{F}$ に関する停止時であるならば，

$$\{\tau\leqq n\}=\bigcup_{k=1}^{n}\{\tau=k\}\in\mathcal{F}_n$$

であり，逆に $\{\tau\leqq n\}\in\mathcal{F}_n$ ならば

$$\{\tau=n\}=\{\tau\leqq n\}-\{\tau\leqq n-1\}\in\mathcal{F}_n$$

となり，τ が $\boldsymbol{F}$ に関する停止時となる．したがって

$$\{\tau\leqq n\}\in\mathcal{F}_n,\quad n=1, 2, \cdots$$

のとき，τ を $\boldsymbol{F}$ に関する停止時であると定義してもよい．このことを用いて，τ_1, τ_2 が $\boldsymbol{F}$ に関する停止時であるとき

$$\tau_1\vee\tau_2=\max(\tau_1, \tau_2),\quad \tau_1\wedge\tau_2=\min(\tau_1, \tau_2),\quad \tau_1+\tau_2$$

も同様であることを証明することができる．すなわち

$$\{\tau_1\vee\tau_2\leqq n\}=\{\tau_1\leqq n\}\cap\{\tau_2\leqq n\}\in\mathcal{F}_n,$$
$$\{\tau_1\wedge\tau_2\leqq n\}=\{\tau_1\leqq n\}\cup\{\tau_2\leqq n\}\in\mathcal{F}_n,$$
$$\{\tau_1+\tau_2\leqq n\}=\bigcup_{k=1}^{n-1}(\{\tau_1\leqq k\}\cap\{\tau_2\leqq n-k\})\in\mathcal{F}_n.$$

定理 5.6 から容易にわかるように，任意の増大情報系 $\boldsymbol{F}=\{\mathcal{F}_1, \mathcal{F}_2, \cdots\}$ はある

確率過程 $X=\{X_1, X_2, \cdots\}$ で生成される．すなわち

$$\mathscr{F}_n = \mathscr{F}[X_1, X_2, \cdots, X_n], \qquad n=1,2,\cdots$$

とかける．これは

$$\mathscr{F}_n = \mathscr{F}[X_{k\wedge n}, k=1,2,\cdots]$$

とかいてもよい．実際右辺は

$$\mathscr{F}[X_1, X_2, \cdots, X_n, X_n, X_n, \cdots] = \mathscr{F}[X_1, X_2, \cdots, X_n]$$

に等しいからである．$\{X_k^{(n)} \equiv X_{k\wedge n}, k=1,2,\cdots\}$ を，$\{X_1, X_2, \cdots\}$ を**時点 n でとめた過程（n 停止過程）**という．この式を背景として，$\boldsymbol{F}$ に関する停止時 τ までに得られる情報を

$$\mathscr{F}_\tau = \mathscr{F}[X_{k\wedge\tau(\omega)}(\omega), k=1,2,\cdots]$$

によって定義したらよいであろう．右辺の定義が可能なためには，$X_{k\wedge\tau(\omega)}(\omega)$ が確率変数となることを確かめる必要があるが，それは

$$\{X_{k\wedge\tau(\omega)}(\omega) \leqq a\} = \bigcup_n \{\tau=n\} \cap \{X_{k\wedge n} \leqq a\} \in \mathscr{D}(P)$$

により明らかである．

上の $\mathscr{F}_\tau$ の定義では，$\mathscr{F}_\tau$ が $\boldsymbol{F}$ を生成する $X=\{X_n\}$（これはいくつもあり得る）に関係する心配があるが，実際は

$$(*) \qquad A \in \mathscr{F}_\tau \Longleftrightarrow A \in \mathscr{F}_\infty, \quad A \cap \{\tau=n\} \in \mathscr{F}_n, \quad n=1,2,\cdots$$

$$\left(\mathscr{F}_\infty = \bigvee_n \mathscr{F}_n\right)$$

$$(\Longleftrightarrow A \cap \{\tau=n\} \in \mathscr{F}_n, \quad n=1,2,\cdots,\infty)$$

$$\left(\{\tau=\infty\} = \Omega - \bigcup_{n<\infty} \{\tau=n\} \in \mathscr{F}_\infty\right)$$

がいえるので，$\mathscr{F}_\tau$ が $\boldsymbol{F}=\{\mathscr{F}_n\}$ だけで定まることがわかる．上の同値性を証明しよう．

$A \in \mathscr{F}_\tau$ ならば，A の指示関数 1_A は

$$1_A = \varphi(X_{k\wedge\tau}, k=1,2,\cdots) \qquad (\varphi : \boldsymbol{R}^\infty \to \{0,1\})$$

とかける．また，$\{\tau=n\} \in \mathscr{F}_n$ により

$$1_{A\cap\{\tau=n\}} = 1_A \cdot 1_{\{\tau=n\}} = \varphi(X_{k\wedge n}, k=1,2,\cdots) \cdot 1_{\{\tau=n\}}$$

が可測 $\mathscr{F}_n$ となり，$A \cap \{\tau=n\} \in \mathscr{F}_n$ $(n=1,2,\cdots,\infty)$ が得られる．したがって $(*)$ の‘$\Longrightarrow$’がなりたつ．

逆に $A \cap \{\tau=n\} \in \mathscr{F}_n$ $(n=1,2,\cdots,\infty)$ ならば，

$$1_{A\cap\{\tau=n\}} = \varphi_n(X_1, X_2, \cdots, X_n) \text{ a.s.} \qquad (\varphi_n : \boldsymbol{R}^n \to \{0,1\}).$$

また $\{\tau=n\} \in \mathscr{F}_n$ により

$$1_{\{\tau=n\}} = \psi_n(X_1, X_2, \cdots, X_n) \text{ a.s.} \qquad (\psi_n : \boldsymbol{R}^n \to \{0,1\}).$$

$\sum_n 1_{\{\tau=n\}}=1$ により，$1_{\{\tau=n\}}$, $n=1,2,\cdots,\infty$ の任意の2個が同時に1になることはない．したがって

$$1_{\{\tau=n\}} = \psi_n'(X_1, X_2, \cdots, X_n) \text{ a.s.}$$

ここに

$$\psi_n'(x_1, x_2, \cdots, x_n) = \prod_{k=1}^{n-1} (1-\psi_k(x_1, x_2, \cdots, x_k)) \cdot \psi_n(x_1, x_2, \cdots, x_n).$$

ただし $\psi_1'(x_1)=\psi_1(x_1)$．ゆえに

$$1_{A\cap\{\tau=n\}} = 1_{\{\tau=n\}} \cdot 1_{A\cap\{\tau=n\}} = \psi_n'(X_1, X_2, \cdots, X_n)\varphi_n(X_1, X_2, \cdots, X_n)$$

となる．さて

$$\theta(x_1, x_2, \cdots) = \sum_{n=1}^{\infty} \psi_n'(x_1, x_2, \cdots, x_n)\varphi_n(x_1, x_2, \cdots, x_n)$$

とおくと，上の注意から

$$\psi_n'(x_1, x_2, \cdots, x_n) = 1 \Longrightarrow \theta(x_1, x_2, \cdots) = \varphi_n(x_1, x_2, \cdots, x_n).$$

上にあらわれた諸式で‘a.s.’がついているものは可算個であるから，Ω から P 零集合を除いておくことにより，例外なしになりたつとしておいてよい．

$$\begin{aligned}
\tau = n &\Longrightarrow \psi_n'(X_1, X_2, \cdots, X_n) = 1 \\
&\Longrightarrow \psi_n'(X_{1\wedge\tau}, X_{2\wedge\tau}, \cdots, X_{n\wedge\tau}) = 1 \qquad (\because\ \tau=n) \\
&\Longrightarrow \theta(X_{1\wedge\tau}, X_{2\wedge\tau}, \cdots, X_{n\wedge\tau}, \cdots) = \varphi_n(X_{1\wedge\tau}, X_{2\wedge\tau}, \cdots, X_{n\wedge\tau}) \\
&\qquad = \varphi_n(X_1, X_2, \cdots, X_n) = 1_{A\cap\{\tau=n\}} = 1_A \qquad (\because\ \tau=n).
\end{aligned}$$

$\Omega = \sum_{n=1}^{\infty} \{\tau=n\}$ により，Ω の上で

$$1_A = \theta(X_{1\wedge\tau}, X_{2\wedge\tau}, \cdots).$$

これは $A \in \mathscr{F}_\tau$ を意味する．したがって $(*)$ の‘$\Longleftarrow$’がなりたつ．

$(*)$ が証明されたから，$\mathscr{F}_\tau$ を

$$\mathscr{F}_\tau = \{A \in \mathscr{F}_\infty \mid A \cap \{\tau=n\} \in \mathscr{F}_n, n=1,2,\cdots\} \qquad \left(\mathscr{F}_\infty = \bigvee_n \mathscr{F}_n\right)$$

と定義してもよいことがわかる．また，この定義は

$$\mathscr{F}_\tau = \{A \in \mathscr{F}_\infty \mid A \cap \{\tau \leqq n\} \in \mathscr{F}_n,\ n=1, 2, \cdots\}$$

とも同等である．

今まで離散時変数の増大情報系について停止時を定義したが，連続時変数のときも類似の論法で進むことができる．ただ連続時変数のためにおこる微妙な点に注意する必要がある．本質的でないことを一々注意しないですむように，$T=[0, \infty)$ のばあいのみを考えよう．実際に重要なのはこのばあいである．

$\boldsymbol{F}=\{\mathscr{F}_t, t \in [0, \infty)\}$ を増大情報系とする．$[0, \infty]$の中の値をとる確率変数 $\tau=\tau(\omega)$ が $\boldsymbol{F}$ に関する停止時(以後簡単に **$\boldsymbol{F}$ 停止時**という)であるとは，

$$\{\tau \leqq t\} \in \mathscr{F}_t, \qquad t \in [0, \infty)$$

がなりたつことであると定義する．これから

$$\{\tau < t\} = \bigcup_n \left\{\tau \leqq t - \frac{1}{n}\right\} \in \mathscr{F}_t.$$

したがって

$$\{\tau = t\} = \{\tau \leqq t\} - \{\tau < t\} \in \mathscr{F}_t$$

がでる．しかし離散時変数のときと違って‘$\{\tau=t\} \in \mathscr{F}_t,\ t \in [0, \infty)$’から $\{\tau \leqq t\} \in \mathscr{F}_t$ はでない．なぜならば

$$\{\tau \leqq t\} = \bigcup_{s \leqq t} \{\tau = s\} \qquad (\text{非可算和})$$

となるからである．したがって‘$\{\tau=t\} \in \mathscr{F}_t,\ t \in [0, \infty)$’で停止時を定義することは適当ではない．

定理 5.8 つぎの諸条件はたがいに同値である．

(i) τ が $\boldsymbol{F}$ 停止時である，すなわち $\{\tau \leqq t\} \in \mathscr{F}_t$,

(ii) $\{\tau < t\} \in \mathscr{F}_t, \quad t \in [0, \infty)$,

(iii) $\{\tau > t\} \in \mathscr{F}_t, \quad t \in [0, \infty)$,

(iv) $\{\tau \geqq t\} \in \mathscr{F}_t, \quad t \in [0, \infty)$.

証明 (i) $\Longleftrightarrow$ (iii), (ii) $\Longleftrightarrow$ (iv) は明らかであり，(i) $\Longrightarrow$ (ii) は上に示したから，(ii) $\Longrightarrow$ (i) をいえばよい．

$$\{\tau \leqq t\} = \bigcap_{n \geqq k} \left\{\tau < t + \frac{1}{n}\right\} \in \mathscr{F}_{t+1/k}, \qquad k = 1, 2, \cdots,$$

ゆえに

$$\{\tau \leqq t\} \in \bigcap_k \mathscr{F}_{t+1/k} = \mathscr{F}_{t+} = \mathscr{F}_t. \qquad \blacksquare$$

$c \in [0, \infty)$ を定数とする．$\tau(\omega) \equiv c$ が $\boldsymbol{F}$ 停止時であることは

$$\{\tau \leqq t\} = \Omega\ (t \geqq c) \quad \text{または} \quad \phi\ (t < c)$$

により明らかである．τ が $\boldsymbol{F}$ 停止時であれば，$\tau + c$ も $\boldsymbol{F}$ 止め時である．

定理 5.9 $\tau_1, \tau_2, \cdots$ を可算個の $\boldsymbol{F}$ 停止時とするとき，

$$\bigvee_n \tau_n = \sup_n \tau_n, \qquad \bigwedge_n \tau_n = \inf_n \tau_n, \qquad \sum_n \tau_n$$

はいずれも $\boldsymbol{F}$ 停止時である．(この事実は離散時変数のときにもなりたつ.)

証明 連続時変数のときにはつぎの事実より明らかである．

$$\left\{\bigvee_n \tau_n \leqq t\right\} = \bigcap_n \{\tau_n \leqq t\} \in \mathscr{F}_t.$$

$$\left\{\bigwedge_n \tau_n < t\right\} = \bigcup_n \{\tau_n < t\} \in \mathscr{F}_t.$$

$$\{\tau_1 + \tau_2 < t\} = \bigcup_{r \in \boldsymbol{Q} \cap [0, t]} (\{\tau_1 < r\} \cap \{\tau_2 < t - r\}) \in \mathscr{F}_t,$$

ゆえに

$$\{\tau_1 + \tau_2 + \cdots + \tau_m < t\} \in \mathscr{F}_t.$$

また明らかに

$$\tau_1 + \tau_2 + \cdots = \sup_m \{\tau_1 + \tau_2 + \cdots + \tau_m\}. \qquad \blacksquare$$

τ を $\boldsymbol{F}=\{\mathscr{F}_t, t \in [0, \infty)\}$ に関する停止時とするとき，離散時変数の時と同様に

$$\mathscr{F}_\tau = \{A \in \mathscr{F}_\infty \mid A \cap \{\tau \leqq t\} \in \mathscr{F}_t,\ t \in [0, \infty)\}, \qquad \mathscr{F}_\infty = \bigvee_t \mathscr{F}_t$$

と定義すると，$\mathscr{F}_\tau$ が閉 σ 加法族(情報)となることは

$$\mathscr{F}_\tau \supset \boldsymbol{2},$$

$$A_n \in \mathscr{F}_\tau \Longrightarrow \left(\bigcup_n A_n\right) \cap \{\tau \leqq t\} = \bigcup_n (A_n \cap \{\tau \leqq t\}) \in \mathscr{F}_t,$$

$$A \in \mathscr{F}_t \Longrightarrow A^c \cap \{\tau \leqq t\} = \{\tau \leqq t\} - A \cap \{\tau \leqq t\} \in \mathscr{F}_t$$

により明らかである．$\mathscr{F}_\tau$ を‘τ までの**情報**’という．

$$\{\tau \leqq s\} \cap \{\tau \leqq t\} = \{\tau \leqq s \wedge t\} \in \mathscr{F}_{s \wedge t} \subset \mathscr{F}_t$$

により $\{\tau \leqq s\} \in \mathscr{F}_\tau$．これは τ が可測 $\mathscr{F}_\tau$ であることを意味する．おおまかにいえば，$\tau(\omega)$ の値は τ までの情報できまるということで，これは停止時の定義の直観的背景であった．

定理 5.10　$\sigma, \tau, \tau_1, \tau_2, \cdots$ はすべて $\boldsymbol{F}=\{\mathcal{F}_t, t\in[0,\infty)\}$ に関する停止時とする.

(i)　$\mathcal{F}_\tau=\{A\in\mathcal{F}_\infty \mid A\cap\{\tau<t\}\in\mathcal{F}_t, t\in[0,\infty)\}$,

(ii)　τ は可測 $\mathcal{F}_\tau$,

(iii)　$\sigma\leqq\tau \Longrightarrow \mathcal{F}_\sigma\subset\mathcal{F}_\tau$,

(iv)　$\mathcal{F}_{\sigma\wedge\tau}=\mathcal{F}_\sigma\wedge\mathcal{F}_\tau,\quad \mathcal{F}_{\wedge\tau_n}=\bigwedge\mathcal{F}_{\tau_n}$,

(v)　$\{\sigma<\tau\}, \{\sigma\leqq\tau\}\in\mathcal{F}_{\sigma\wedge\tau}$,

(vi)　$\mathcal{F}_{\sigma\vee\tau}=\mathcal{F}_\sigma\vee\mathcal{F}_\tau,\quad \mathcal{F}_{\vee\tau_n}\supset\bigvee_n\mathcal{F}_{\tau_n}$.

証明　(i)　下の式から容易に証明される.

$$A\cap\{\tau<t\}=\bigcup_n A\cap\left\{\tau\leqq t-\frac{1}{n}\right\},$$

$$A\cap\{\tau\leqq t\}=\bigcap_{n\geqq k}A\cap\left\{\tau<t+\frac{1}{n}\right\},\qquad k=1,2,\cdots.$$

(ii)　上で証明した.

(iii)　$\sigma\leqq\tau$ ならば, $\{\sigma\leqq t\}\supset\{\tau\leqq t\}$. したがって

$$A\in\mathcal{F}_\sigma \Longrightarrow A\cap\{\tau\leqq t\}=(A\cap\{\sigma\leqq t\})\cap\{\tau\leqq t\}\in\mathcal{F}_t$$
$$\Longrightarrow A\in\mathcal{F}_\tau.$$

(iv)　上に示した (iii) により, $\mathcal{F}_{\wedge\tau_n}\subset\bigwedge\mathcal{F}_{\tau_n}$.

$$A\in\bigwedge\mathcal{F}_{\tau_n} \Longrightarrow A\in\mathcal{F}_{\tau_n},\qquad n=1,2,\cdots$$
$$\Longrightarrow A\cap\{\tau_n<t\}\in\mathcal{F}_t,\qquad n=1,2,\cdots$$
$$\Longrightarrow A\cap\{\wedge\tau_n<t\}=\bigcup_n(A\cap\{\tau_n<t\})\in\mathcal{F}_t$$
$$\Longrightarrow A\in\mathcal{F}_{\wedge\tau_n}.$$

ゆえに $\bigwedge\mathcal{F}_{\tau_n}\subset\mathcal{F}_{\wedge\tau_n}$. $\mathcal{F}_{\sigma\wedge\tau}=\mathcal{F}_\sigma\wedge\mathcal{F}_\tau$ も同様.

(v)
$$\{\sigma<\tau\}\cap\{\sigma<t\}=\bigcup_{r<t}\{\sigma<r\}\cap\{\tau>r\}\cap\{\sigma<t\}\qquad(r\in\boldsymbol{Q})$$
$$\in\mathcal{F}_t.$$

したがって $\{\sigma<\tau\}\in\mathcal{F}_\sigma$. 同様に $\{\sigma<\tau\}\in\mathcal{F}_\tau$. したがって, (iv) により

$$\{\sigma<\tau\}\in\mathcal{F}_\sigma\wedge\mathcal{F}_\tau=\mathcal{F}_{\sigma\wedge\tau}.$$

また上の関係と (iii) を用いて

$$\{\sigma\leqq\tau\}=\bigcap_{n\geqq k}\left\{\sigma<\tau+\frac{1}{n}\right\}\in\mathcal{F}_{\sigma\wedge(\tau+1/k)},\qquad k=1,2,\cdots,$$

ゆえに (iv) により

$$\{\sigma\leqq\tau\}\in\bigwedge_k\mathcal{F}_{\sigma\wedge(\tau+1/k)}=\mathcal{F}_{\wedge(\sigma\wedge(\tau+1/k))}=\mathcal{F}_{\sigma\wedge\tau}.$$

(vi)　(iii) により第2式は明らか. 同様に $\mathcal{F}_{\sigma\vee\tau}\supset\mathcal{F}_\sigma\vee\mathcal{F}_\tau$. 逆の包含関係を示すために, 任意の $A\in\mathcal{F}_{\sigma\vee\tau}$ が $\mathcal{F}_\sigma\vee\mathcal{F}_\tau$ に属することをいえばよい. それには

$$A=(A\cap\{\sigma\leqq\tau\})\cup(A\cap\{\tau\leqq\sigma\})$$

により,

$$A\cap\{\sigma\leqq\tau\}\in\mathcal{F}_\tau\ (\subset\mathcal{F}_\sigma\vee\mathcal{F}_\tau),\qquad A\cap\{\tau\leqq\sigma\}\in\mathcal{F}_\sigma\ (\subset\mathcal{F}_\sigma\vee\mathcal{F}_\tau)$$

をいえば十分である. $\{\sigma\leqq\tau\}$ ならば $\sigma\vee\tau=\tau$ であるから

$$A\cap\{\sigma\leqq\tau\}\cap\{\tau\leqq t\}=A\cap\{\sigma\leqq\tau\}\cap\{\tau\leqq t\}\cap\{\sigma\vee\tau\leqq t\},$$
$$A\in\mathcal{F}_{\sigma\vee\tau}\quad により\quad A\cap\{\sigma\vee\tau\leqq t\}\in\mathcal{F}_t,$$
$$\{\sigma\leqq\tau\}\in\mathcal{F}_{\sigma\wedge\tau}\subset\mathcal{F}_\tau\quad により\quad \{\sigma\leqq\tau\}\cap\{\tau\leqq t\}\in\mathcal{F}_t,$$

したがって

$$A\cap\{\sigma\leqq\tau\}\cap\{\tau\leqq t\}\in\mathcal{F}_t.$$

これは $A\cap\{\sigma\leqq\tau\}\in\mathcal{F}_\tau$ を意味する. 同様に $A\cap\{\tau\leqq\sigma\}\in\mathcal{F}_\sigma$ が得られる. ▌

例題 5.4　$\boldsymbol{F}=\{\mathcal{F}_t, t\in[0,\infty)\}$ を増大情報系とする.

(i)　$\{\tau_n\}$ が $\boldsymbol{F}$ 停止時の減少列ならば, $\tau=\lim\tau_n$ も $\boldsymbol{F}$ 停止時で, $\mathcal{F}_\tau=\bigwedge_n\mathcal{F}_{\tau_n}$ となることを示せ.

［ヒント］　$\tau=\wedge\tau_n$ に注意し, 定理 5.10 (iv) を用いる.

(ii)　任意の $\boldsymbol{F}$ 停止時 τ は離散値をとる $\boldsymbol{F}$ 停止時の減少列の極限であることを示せ.

［ヒント］　$\tau_n=2^{-n}([2^n\tau]+1)$ ([　] は整数部分) とおくと, τ_n は $2^{-n}k$ $(k=0,1,2,\cdots,\infty)$ の値をとる $\boldsymbol{F}$ 停止時で, $\tau_n\downarrow\tau$.

(iii)　$X=\{X_t, t\in[0,\infty)\}$ を確率過程とする. 任意の集合 $E\subset\boldsymbol{R}^1$ に対し

$$\sigma_E(\omega)=\inf\{t\geqq0 \mid X_t(\omega)\in E\}$$

を X の E への**到達時**(hitting time) という. X が $\boldsymbol{F}$ に適合した C 過程で, E が開集合また閉集合のときには, σ_E は $\boldsymbol{F}$ 停止時であることを示せ.

［ヒント］　E が開集合ならば

$$\{\sigma_E<t\}=\bigcup_{r<t}\{X_r\in E\}\qquad(r\ は有理数)$$

に注意すれば, $\{\sigma_E<t\}\in\mathcal{F}_t$ がでる. E が閉集合ならば, $E=\bigcap E_n$ (E_n は開,

$E_n \supset \bar{E}_{n+1}$）とあらわせるが，$\sigma_E = \bigvee_n \sigma_{E_n}$ となることが，$X_t(\omega)$ が t に関して連続なことからわかる．

§5.5 離散時変数のマルチンゲール

マルチンゲールは公平な賭けという考えから生れた概念であるが，その一般論は J. L. Doob によってつくられた．

X を可積分な実確率変数，$\boldsymbol{F}$ を閉 σ 加法族（情報）とする．$\mathscr{F}$ のもとにおける条件付平均値 $E_{\mathscr{F}}(X)$ (§3.8) を $E(X|\mathscr{F})$ とかくことも多い．とくに $\mathscr{F}=\mathscr{F}[X_1, X_2, \cdots, X_n]$ のときには，$E(X|X_1, X_2, \cdots, X_n)$ とかくことにする．また

$$L_{\mathscr{F}}{}^p = \{X \in L^p \,|\, X \text{ は可測 } \mathscr{F}\}$$

を $L^p(\mathscr{F})$ であらわすこともある．

$X=\{X_0, X_1, X_2, \cdots\}$ を実確率変数列（離散時変数）の確率過程とし，$\boldsymbol{F}=\{\mathscr{F}_0, \mathscr{F}_1, \mathscr{F}_2, \cdots\}$ を増大情報系とする．（以後時変数は 0, 1, 2, … をとる．）もし

(M. 1) X が $\boldsymbol{F}$ に適合，

(M. 2) $X_n \in L^1$（すなわち $E|X_n|<\infty$），$n=0, 1, 2, \cdots$,

(M. 3) $E(X_{n+1}|\mathscr{F}_n) = X_n$ a.s., $n=0, 1, 2, \cdots$

がなりたつならば，X は '**$\boldsymbol{F}$ に関してマルチンゲール** (martingale)'，または簡単に **$\boldsymbol{F}$ マルチンゲール**という．とくに X が $\boldsymbol{F}[X]$（X で生成される増大情報系）に関してマルチンゲールのとき，単に**マルチンゲール**という．以後条件付確率に関する等式や不等式で誤解の心配がない限り 'a.s.' を省略する．

(M. 1), (M. 2) の 2 条件はつぎの 1 条件と同値である：

$$X_n \in L^1(\mathscr{F}_n), \quad n = 0, 1, 2, \cdots.$$

また (M. 3) はつぎの条件と同値である．

$$E(X_{n+1}, A) = E(X_n, A), \quad A \in \mathscr{F}_n, \quad n = 0, 1, 2, \cdots.$$

これは

$$E(X_{n+1}\cdot Y) = E(X_n \cdot Y), \quad Y \in L^\infty(\mathscr{F}_n), \quad n = 0, 1, 2, \cdots$$

といってもよい．

X は当然 $\boldsymbol{F}[X]$ に適合しているから，X がマルチンゲールであるというのは，

$$E|X_n| < \infty, \quad E(X_{n+1}|X_1, X_2, \cdots, X_n) = X_n, \quad n = 0, 1, 2, \cdots$$

を意味する．これは多くの教科書に見られる定義である．

Y を任意の可積分実確率変数，$\boldsymbol{F}=\{\mathscr{F}_0, \mathscr{F}_1, \mathscr{F}_2, \cdots\}$ を増大情報系とするとき，

$$X_n = E(Y|\mathscr{F}_n), \quad n = 0, 1, 2, \cdots$$

は $\boldsymbol{F}$ マルチンゲールである．実際 $X_n \in L^1(\mathscr{F}_n)$ は条件付確率の定義により明らかであり，$\mathscr{F}_{n+1} \supset \mathscr{F}_n$ により

$$E(X_{n+1}|\mathscr{F}_n) = E(E(Y|\mathscr{F}_{n+1})|\mathscr{F}_n) = E(Y|\mathscr{F}_n) = X_n.$$

$X_0, X_1, \cdots$ を独立な可積分実確率変数の列とし，

$$S_n = X_0 + X_1 + \cdots + X_n, \quad \mathscr{F}_n = \mathscr{F}[X_0, X_1, \cdots, X_n], \quad n = 0, 1, 2, \cdots$$

とおく．明らかに $S_n \in L^1(\mathscr{F}_n)$．したがって

$$E(S_n|\mathscr{F}_n) = S_n, \quad n = 0, 1, 2, \cdots.$$

$\{X_n\}$ の独立性により，X_{n+1} は $\mathscr{F}_n$ と独立であるから，

$$E(X_{n+1}|\mathscr{F}_n) = E(X_{n+1}), \quad n = 0, 1, 2, \cdots.$$

したがって

$$E(S_{n+1}|\mathscr{F}_n) = E(S_n|\mathscr{F}_n) + E(X_{n+1}|\mathscr{F}_n) = S_n + E(X_{n+1}).$$

これから，$\{S_n\}$ が $\{\mathscr{F}_n\}$ マルチンゲールであるための必要十分条件は

$$E(X_{n+1}) = 0, \quad n = 0, 1, 2, \cdots$$

であることがわかる．

$\boldsymbol{F}$ マルチンゲール性は（非斉次）1 次結合に関して不変である．すなわち，$X=\{X_n\}$，$Y=\{Y_n\}$ が $\boldsymbol{F}$ マルチンゲールのときには $Z_n = aX_n + bY_n + c$（a, b, c は実定数）も $\boldsymbol{F}$ マルチンゲールである．しかしマルチンゲール性に関してはこのことはいえない．それは，$\boldsymbol{F}[X]$ と $\boldsymbol{F}[Y]$ とは必ずしも一致しないからである．

$\boldsymbol{F}$ マルチンゲールは当然マルチンゲールである．実際 $X=\{X_n\}$ が $\boldsymbol{F}=\{\mathscr{F}_n\}$ に関してマルチンゲールであれば，

$$\mathscr{F}[X_1, X_2, \cdots, X_n] \subset \mathscr{F}_n$$

であるから，

$$\begin{aligned} E(X_{n+1}|\mathscr{F}[X_1, X_2, \cdots, X_n]) &= E(E(X_{n+1}|\mathscr{F}_n)|\mathscr{F}[X_1, X_2, \cdots, X_n]) \\ &= E(X_n|\mathscr{F}[X_1, X_2, \cdots, X_n]) \\ &= X_n \end{aligned}$$

が得られる．このことから，マルチンゲールであるとは，適当な増大情報系に関してマルチンゲールであることと定義することができる．

$X=\{X_n\}$ が $\boldsymbol{F}=\{\mathscr{F}_n\}$ に関してマルチンゲールであるならば，$\mathscr{F}_{n+1} \supset \mathscr{F}_n$ に

より

$$E(X_{n+2}\,|\,\mathcal{F}_n) = E(E(X_{n+2}\,|\,\mathcal{F}_{n+1})\,|\,\mathcal{F}_n) = E(X_{n+1}\,|\,\mathcal{F}_n) = X_n.$$

これをくりかえして

$$E(X_{n+k}\,|\,\mathcal{F}_n) = X_n, \qquad k = 1, 2, \cdots.$$

この式は $k=0$ のときにもなりたつことは，X_n が可測 $\mathcal{F}_n$ であることから明らかである．上の式の両辺の平均値をとって

$$E(X_n) = \text{一定}$$

であることがわかる．

$\boldsymbol{F}$ マルチンゲールの定義の条件の (M. 3) を

$$(\mathrm{M}^+.\,3) \qquad E(X_{n+1}\,|\,\mathcal{F}_n) \geqq X_n$$

でおきかえて，**$\boldsymbol{F}$ 劣マルチンゲール**を定義する．同様に (M. 3) を

$$(\mathrm{M}^-.\,3) \qquad E(X_{n+1}\,|\,\mathcal{F}_n) \leqq X_n$$

でおきかえて **$\boldsymbol{F}$ 優マルチンゲール**を定義しよう．単なる‘**劣マルチンゲール**’‘**優マルチンゲール**’の定義は‘マルチンゲール’のばあいと同様の考えで行う．

$X=\{X_n\}$ が $\boldsymbol{F}=\{\mathcal{F}_n\}$ に関して劣(優)マルチンゲールならば，$-X=\{-X_n\}$ は $\boldsymbol{F}$ に関して優(劣)マルチンゲールである．X が $\boldsymbol{F}$ マルチンゲールであることは“X が $\boldsymbol{F}$ 優マルチンゲールであり，かつ $\boldsymbol{F}$ 劣マルチンゲールである”ことと同値である．

マルチンゲールに関してのべた上述の事実は，等号‘$=$’を不等号‘$\geqq$’(または‘$\leqq$’)でおきかえることにより，劣(または優)マルチンゲールに対してもなりたつ．たとえば，$X=\{X_n\}$ が $\boldsymbol{F}$ 劣マルチンゲールならば，

$$E(X_{n+k}\,|\,\mathcal{F}_n) \geqq X_n, \qquad k = 1, 2, \cdots,$$

$$E(X_n)\ \text{は}\ n\ \text{とともに増大する}.$$

$X=\{X_n\}$，$Y=\{Y_n\}$ が $\boldsymbol{F}$ 劣マルチンゲールならば，

$$Z_n = aX_n + bY_n + c \qquad (a, b\ \text{は正定数},\ c\ \text{は実定数})$$

も $\boldsymbol{F}$ 劣マルチンゲールである．

$A_n \in L^1(\mathcal{F}_n)$ が n とともに増大するならば，$A=\{A_n\}$ が $\boldsymbol{F}=\{\mathcal{F}_n\}$ に関して劣マルチンゲールとなることは自明である．したがってこのような A と $\boldsymbol{F}$ マルチンゲール $M=\{M_n\}$ との和 $M+A=\{M_n+A_n\}$ もまた $\boldsymbol{F}$ 劣マルチンゲールである．実は任意の $\boldsymbol{F}$ 劣マルチンゲールはこのような和としてあらわされる．

定理 5.11 (Doob の分解定理)　任意の $\{\mathcal{F}_n\}$ 劣マルチンゲール $\{X_n\}$ は

$$X_n = M_n + A_n,$$

$$M = \{M_n\}\ \text{は}\ \boldsymbol{F}\ \text{マルチンゲール},$$

$$A = \{A_n\}\ \text{は増大列で},\quad A_0 = 0,\quad A_n \in L^1(\mathcal{F}_{n-1}), \qquad n = 1, 2, \cdots$$

の形に分解され，しかもこの分解は一意である．

証明　もし上の分解が可能ならば，$A_{n+1}-A_n \in L^1(\mathcal{F}_n)$ であるから，

$$\begin{aligned} A_{n+1}-A_n &= E(A_{n+1}-A_n\,|\,\mathcal{F}_n) \\ &= E(X_{n+1}-X_n\,|\,\mathcal{F}_n) - E(M_{n+1}-M_n\,|\,\mathcal{F}_n) \\ &= E(X_{n+1}-X_n\,|\,\mathcal{F}_n), \end{aligned}$$

$A_0=0$ であるから，

$$A_n = \sum_{k=0}^{n-1} E(X_{k+1}-X_k\,|\,\mathcal{F}_k), \qquad M_n = X_n - A_n$$

となるはずである．これから分解の一意性がわかる．

分解の可能性をいうには，上のように A_n, M_n を定めて，上の分解に関する条件がみたされることを検証すればよい．上の A_n, M_n の定義式により，$A_n, M_{n-1} \in L^1(\mathcal{F}_{n-1})$ は明らかであり，また $\boldsymbol{F}$ 劣マルチンゲールの定義により，

$$E(X_{k+1}-X_k\,|\,\mathcal{F}_k) = E(X_{k+1}\,|\,\mathcal{F}_k) - X_k \geqq 0$$

であるから，$\{A_n\}$ の増大性は自明である．A_n, M_n の定義から

$$M_{n+1}-M_n = (X_{n+1}-X_n) - (A_{n+1}-A_n)$$

$$A_{n+1}-A_n = E(X_{n+1}-X_n\,|\,\mathcal{F}_n)$$

$$E(A_{n+1}-A_n\,|\,\mathcal{F}_n) = E(X_{n+1}-X_n\,|\,\mathcal{F}_n)$$

$$E(M_{n+1}-M_n\,|\,\mathcal{F}_n) = 0 \quad \text{すなわち} \quad E(M_{n+1}\,|\,\mathcal{F}_n) = M_n$$

が得られるから，$\{M_n\}$ は $\boldsymbol{F}$ マルチンゲールである．▌

上の分解定理で

$$X_n, M_n \in L^1(\mathcal{F}_n)$$

であるが，定理の第3条件では

(i)　$A_n \in L^1(\mathcal{F}_{n-1})$

となっている．これをなぜ

(ii)　$A_n \in L^1(\mathcal{F}_n)$

としてはいけないのかという疑問がおこるであろう．(ii) は (i) より弱い条件で

あるから，(ii)をとったとしても分解の存在は当然なりたつが，分解の一意性の方は必ずしも保証されない．例えば

$$\Delta_1, \Delta_2, \cdots \text{ は独立な可積分正値確率変数,}$$

$$X_n = \sum_{k=1}^{n} \Delta_k \qquad (X_0=0),$$

$$\mathcal{F}_n = \mathcal{F}[\Delta_1, \Delta_2, \cdots, \Delta_n] \qquad (\text{ただし } \mathcal{F}_0 = \mathbf{2})$$

とすれば，$\{X_n\}$ は $\{\mathcal{F}_n\}$ に関して劣マルチンゲールである．定理において(i)のかわりに(ii)をとれば

$$X_n = 0+X_n,$$

$$X_n = (X_n - E(X_n)) + E(X_n)$$

がともにDoobの分解となるが，上の定理のように(i)をとれば，第2式だけがDoobの分解となるのである．

定理 5.12 f が区間 I (開，閉，半開いずれでもよい，また無限区間でもよい)の上で凸関数とし，

$$X_n(\omega) \in I \text{ a.s.}, \qquad n = 0, 1, 2, \cdots$$

とする．さらに $f(X_n)$，$n=0, 1, 2, \cdots$ はすべて可積分とする．

(i) $\{X_n\}$ が $\{\mathcal{F}_n\}$ マルチンゲールならば，$\{f(X_n)\}$ は $\{\mathcal{F}_n\}$ 劣マルチンゲールである．

(ii) $\{X_n\}$ が $\{\mathcal{F}_n\}$ 劣マルチンゲールで，しかも f が増加関数ならば，$\{f(X_n)\}$ は $\{\mathcal{F}_n\}$ 劣マルチンゲールである．

証明 Jensenの不等式(§3.8)を用いて

$$E(f(X_{n+1}) \mid \mathcal{F}_n) \geqq f(E(X_{n+1} \mid \mathcal{F}_n)).$$

(i)のばあいには $E(X_{n+1} \mid \mathcal{F}_n) = X_n$ となるから

$$E(f(X_{n+1}) \mid \mathcal{F}_n) \geqq f(X_n)$$

が得られる．(ii)のばあいには，仮定により

$$E(X_{n+1} \mid \mathcal{F}_n) \geqq X_n, \qquad \text{したがって} \quad f(E(X_{n+1} \mid \mathcal{F}_n)) \geqq f(X_n)$$

となるから，やはり上の不等式が得られる．▌

上の定理によりつぎのことがわかる：

"$\{X_n\}$ が $\{\mathcal{F}_n\}$ マルチンゲールならば，$\{|X_n|\}$ は $\{\mathcal{F}_n\}$ 劣マルチンゲールであり"，

"$\{X_n\}$ が $\{\mathcal{F}_n\}$ 劣マルチンゲールならば，$\{X_n^+ = X_n \vee 0\}$ は $\{\mathcal{F}_n\}$ 劣マルチンゲールである"．

$\{X_n\}$ が $\{\mathcal{F}_n\}$ マルチンゲールのとき，$n \leqq m$ に対して

$$E(X_m \mid \mathcal{F}_n) = X_n \text{ a.s.}$$

がなりたつことはすでに証明した．m, n を停止時 $\tau(\omega), \sigma(\omega)$ でおきかえたばあいにも同様の事実がなりたつ．

定理 5.13 $\{X_n\}$ が $\{\mathcal{F}_n\}$ マルチンゲールとする．σ, τ が $\{\mathcal{F}_n\}$ 停止時で，$\sigma \leqq \tau \leqq N$ (定数) a.s. ならば

$$E(X_\tau \mid \mathcal{F}_\sigma) = X_\sigma \text{ a.s.}, \qquad \text{ゆえに} \quad E(X_\tau) = E(X_\sigma).$$

また $\{X_n\}$ が $\{\mathcal{F}_n\}$ 劣マルチンゲールのときには

$$E(X_\tau \mid \mathcal{F}_\sigma) \geqq X_\sigma \text{ a.s.}, \qquad \text{ゆえに} \quad E(X_\tau) \geqq E(X_\sigma).$$

証明 マルチンゲールのときについて証明する．

$$E(|X_\tau|) = \sum_{k=0}^{N} E(|X_k|, \{\tau=k\}) \leqq \sum_{k=0}^{N} E(|X_k|) < \infty$$

であるから，問題の条件付平均値は考えられる．$E(X_\tau \mid \mathcal{F}_\sigma)$ は可測 $\mathcal{F}_\sigma$．また

$$\{X_\sigma < a\} \cap \{\sigma=n\} = \{X_n < a\} \cap \{\sigma=n\} \in \mathcal{F}_n$$

であるから，X_σ も可測 $\mathcal{F}_\sigma$．したがって問題の等式を示すには，

$$E(E(X_\tau \mid \mathcal{F}_\sigma), A) = E(X_\sigma, A) \quad \text{すなわち} \quad E(X_\tau, A) = E(X_\sigma, A)$$

が任意の $A \in \mathcal{F}_\sigma$ に対しなりたつことをいえばよい．

まず $\tau=\sigma$ または $\sigma+1$ のときには定理がなりたつことを示そう．

$$\{\tau=\sigma\} \in \mathcal{F}_\sigma \quad (\because \{\tau=\sigma\} \cap \{\sigma=n\} = \{\tau=\sigma=n\} \in \mathcal{F}_n),$$

$$\{\tau=\sigma+1\} = \{\tau=\sigma\}^c \in \mathcal{F}_\sigma$$

であるから，仮定 $A \in \mathcal{F}_\sigma$ により

$$A \cap \{\tau=\sigma\} \in \mathcal{F}_\sigma, \qquad A \cap \{\tau=\sigma+1\} \in \mathcal{F}_\sigma,$$

したがって

$$A \cap \{\tau=\sigma\} \cap \{\sigma=k\} \in \mathcal{F}_k, \qquad A \cap \{\tau=\sigma+1\} \cap \{\sigma=k\} \in \mathcal{F}_k$$

となる．さて

$$E(X_\tau, A) = E(X_\tau, A \cap \{\tau=\sigma\}) + E(X_\tau, A \cap \{\tau=\sigma+1\})$$

$$= \sum_{k=0}^{N} E(X_\tau, A \cap \{\tau=\sigma\} \cap \{\sigma=k\})$$

$$+\sum_{k=0}^{N} E(X_\tau, A\cap\{\tau=\sigma+1\}\cap\{\sigma=k\})$$
$$=\sum_{k=0}^{N} E(X_k, A\cap\{\tau=\sigma\}\cap\{\sigma=k\})$$
$$+\sum_{k=0}^{N} E(X_{k+1}, A\cap\{\tau=\sigma+1\}\cap\{\sigma=k\}).$$

$\{X_n\}$ は $\{\mathcal{F}_n\}$ マルチンゲールであるから，上に注意したことから，第2項の X_{k+1} は X_k でおきかえることができて，

$$=\sum_{k=0}^{N} E(X_k, A\cap\{\tau=\sigma\}\cap\{\sigma=k\})+\sum_{k=0}^{N} E(X_k, A\cap\{\tau=\sigma+1\}\cap\{\sigma=k\})$$
$$=E(X_\sigma, A\cap\{\tau=\sigma\})+E(X_\sigma, A\cap\{\tau=\sigma+1\})$$
$$=E(X_\sigma, A).$$

これで $\tau=\sigma$ または $\sigma+1$ のときには定理がなりたつ．

一般のばあいを示すには

$$\sigma_0=\sigma, \qquad \sigma_k=\tau\wedge(\sigma+k),\ k=1,2,\cdots$$

とおくと，$\sigma_{k+1}=\sigma_k$ または σ_k+1 であるから，上に示したことから

$$E(X_{\sigma_{k+1}}\,|\,\mathcal{F}_{\sigma_k})=X_{\sigma_k}, \qquad k=0,1,2,\cdots$$

がでる．これは $\{X_n'=X_{\sigma_n}\}$ が $\{\mathcal{F}_n'=\mathcal{F}_{\sigma_n}\}$ に関してマルチンゲールであることを示している．したがって

$$E(X_{\sigma_N}\,|\,\mathcal{F}_{\sigma_0})=X_{\sigma_0}$$

が得られる．$\sigma\leqq\tau\leqq N$ により $\sigma_N=\tau\wedge(\sigma+N)=\tau$ であり，σ_0 の定義により $\sigma_0=\sigma$ であるから，上の式は定理の結論を示している．▌

$X=\{X_n\}$ を $\boldsymbol{F}=\{\mathcal{F}_n\}$ に適合する確率過程とする．$\boldsymbol{F}$ 停止時 σ に対して

$$X^\sigma=\{X_n^\sigma\equiv X_{n\wedge\sigma}, n=0,1,2,\cdots\}$$

を X の σ **停止過程**という．また

$$\boldsymbol{F}^\sigma=\{\mathcal{F}_n^\sigma\equiv\mathcal{F}_{n\wedge\sigma}, n=0,1,2,\cdots\}$$

を $\boldsymbol{F}$ の σ **停止増大情報系**という．この操作は観察を時点 σ で停止することに対応する．σ が定数でなく確率変数であるという点を強調するため**任意停止** (optional stopping) ということもある．明らかに X^σ は $\boldsymbol{F}^\sigma$ に適合する．

$\boldsymbol{F}$ 停止時の増大列 $\Sigma=\{\sigma_0,\sigma_1,\sigma_2,\cdots\}$ $(\sigma_0\leqq\sigma_1\leqq\sigma_2\leqq\cdots)$ に対して，

$$X^\Sigma=\{X_n^\Sigma\equiv X_{\sigma_n}, n=0,1,2,\cdots\},$$

$$\boldsymbol{F}^\Sigma=\{\mathcal{F}_n^\Sigma\equiv\mathcal{F}_{\sigma_n}, n=0,1,2,\cdots\}$$

をそれぞれ X の Σ **抽出過程**，$\boldsymbol{F}$ の Σ **抽出増大系**という．この操作は時点 $\sigma_0,\sigma_1,\sigma_2,\cdots$ を抽出して，その時点においてのみ観察することを意味する．抽出時点が確率変数であるから，**任意抽出** (optional sampling) ということもある．明らかに X^Σ は $\boldsymbol{F}^\Sigma$ に適合する．

Σ として $\{\sigma_n\equiv\sigma\wedge n\}$ をとると

$$X^\Sigma=X^\sigma, \qquad \boldsymbol{F}^\Sigma=\boldsymbol{F}^\sigma$$

となるから，任意停止は任意抽出の特別のばあいである．

マルチンゲール性は任意停止や任意抽出で不変である．すなわち

(i) **任意停止定理** X が $\boldsymbol{F}$ マルチンゲールならば，X^σ は $\boldsymbol{F}^\sigma$ マルチンゲールである．優マルチンゲール，劣マルチンゲールについても同様，

(ii) **任意抽出定理** X が $\boldsymbol{F}$ マルチンゲールで，

$$\sigma_n\leqq N_n, \quad n=0,1,\cdots \qquad (N_n \text{ は } n \text{ に関係する定数})$$

ならば，X^Σ は $\boldsymbol{F}^\Sigma$ マルチンゲールである．優マルチンゲール，劣マルチンゲールについても同様

がなりたつ．実際 (i) は (ii) の特別のばあいである．(ii) を示すには前定理を利用すればよい．

定理 5.14 (**Doob の不等式**) $\{X_n\}$ が $\{\mathcal{F}_n\}$ 劣マルチンゲールで，$X_n\geqq0$ a.s. $(n=0,1,\cdots)$ ならば

$$P\left\{\max_{k=0}^{N} X_k\geqq a\right\}\leqq\frac{1}{a}E(X_N), \qquad a>0.$$

証明 上の式の中の最大値を Y_N とする．

$$\sigma=\begin{cases}\inf\{k\leqq N\,|\,X_k\geqq a\} & (Y_N\geqq a \text{ のとき}),\\ N & (Y_N<a \text{ のとき})\end{cases}$$

とすると，$\{Y_n\}$ も $\{\mathcal{F}_n\}$ に適合するから，$k<N$ のときには

$$\{\sigma\leqq k\}=\{Y_k\geqq a\}\in\mathcal{F}_k,$$

$k\geqq N$ のときには

$$\{\sigma\leqq k\}=\Omega\in\mathcal{F}_k$$

となる．したがって σ は $\{\mathcal{F}_n\}$ 停止時である．$\sigma\leqq N$ により

$$E(X_N)\geqq E(X_\sigma) \qquad (\text{定理 5.13 により}).$$

$X_\sigma \geqq 0$ a.s. で，とくに $\{Y_N \geqq a\}$ の上では，σ の定義により $X_\sigma \geqq a$ であるから，

$$E(X_N) \geqq E(X_\sigma) \geqq aP\{Y_N \geqq a\}. \qquad \blacksquare$$

$X=\{X_n\}$ が $\boldsymbol{F}=\{\mathcal{F}_n\}$ に適合する確率過程とする．実数の区間 $[a,b]$ $(a<b)$ に対して

$$X_n \leqq a < X_{n+1}, X_{n+2}, \cdots, X_{m-1} < b \leqq X_m$$

がなりたつとき，X は時区間 $[n,m]$ で $[a,b]$ を上向きに渡るといい，$[n,m]$ を $[a,b]$ に対する X の**上渡時区間**という．上渡時区間はいくつもあるが，たがいに離れている．$[0,N]$ に含まれる上渡時区間の数を $[a,b]$ に対する，時点 N までの**上渡回数**といい，β_N であらわす．β_N は当然 N をこえない．

$$\sigma_1 = \inf\{k \mid X_k \leqq a\},$$
$$\tau_1 = \inf\{k>\sigma_1 \mid X_k \geqq b\},$$
$$\sigma_2 = \inf\{k>\tau_1 \mid X_k \leqq a\},$$
$$\tau_2 = \inf\{k>\sigma_2 \mid X_k \geqq b\},$$
$$\cdots\cdots\cdots\cdots\cdots$$

とする．ただし $\inf\phi$ は ∞ とする．明らかにつぎの3通りの可能性がある．

$$\sigma_1 < \tau_1 < \sigma_2 < \tau_2 < \cdots,$$
$$\sigma_1 < \tau_1 < \sigma_2 < \tau_2 < \cdots < \sigma_m < \tau_m = \tau_{m+1} = \cdots = \infty,$$
$$\sigma_1 < \tau_1 < \sigma_2 < \tau_2 < \cdots < \sigma_{m-1} < \tau_{m-1} < \sigma_m = \sigma_{m+1} = \cdots = \infty.$$

いずれのばあいにも $\sigma_i<\tau_i$ ならば $[\sigma_i,\tau_i]$ はただ1個の上渡時区間を含み，逆に任意の上渡時区間はただ1個の $[\sigma_i,\tau_i]$ $(\sigma_i<\tau_i)$ に含まれる．したがって

$$\beta_N = \sup\{i \mid \tau_i \leqq N\}$$

となる．

$\beta_N=k$ $(=0,1,2,\cdots,N)$ となるのは，

(1) $\sigma_1 < \tau_1 < \sigma_2 < \tau_2 < \cdots < \sigma_k < \tau_k \leqq N < \sigma_{k+1} \leqq \tau_{k+1} \leqq \cdots$

または

(2) $\sigma_1 < \tau_1 < \sigma_2 < \tau_2 < \cdots < \sigma_k < \tau_k < \sigma_{k+1} \leqq N < \tau_{k+1} \leqq \sigma_{k+2} \leqq \cdots$

のばあいである．

さて

$$Z_N = \sum_{i=1}^{N}(X_{\sigma_{i+1}\wedge N} - X_{\tau_i\wedge N})$$

とおく．$\beta_N=k$ $(=0,1,2,\cdots,N)$ とせよ．(1), (2) いずれのばあいにも，$i \geqq k+1$ のときには $\sigma_{i+1}, \tau_i \geqq N$ となるから，上の和の $(k+1)$ 項以後は消えて

$$Z_N = \sum_{i=1}^{k}(X_{\sigma_{i+1}\wedge N} - X_{\tau_i\wedge N})$$

となる．(1) のときには

$$Z_N = \sum_{i=1}^{k-1}(X_{\sigma_{i+1}} - X_{\tau_i}) + (X_N - X_{\tau_k})$$
$$= \sum_{i=1}^{k-1}(X_{\sigma_{i+1}} - X_{\tau_i}) + (a - X_{\tau_k}) + (X_N - a),$$

τ_i, σ_i の定義により

$$\leqq (k-1)(a-b) + (a-b) + (X_N - a)$$
$$\leqq k(a-b) + (X_N-a)^+,$$

(2) のときには

$$Z_N = \sum_{i=1}^{k}(X_{\sigma_{i+1}} - X_{\tau_i}) \leqq k(a-b) \leqq k(a-b) + (X_N-a)^+,$$

いずれのばあいにも

$$\text{(a)} \qquad Z_N \leqq \beta_N(a-b) + (X_N-a)^+$$

となる．この不等式から，つぎの重要な結果が導かれる．

定理 5.15 (Doob の上渡回数定理) $X=\{X_n\}$ が $\boldsymbol{F}=\{\mathcal{F}_n\}$ に関して劣マルチンゲールならば，

$$E(\beta_N) \leqq \frac{E((X_N-a)^+)}{b-a}.$$

証明 上に得た不等式 (a) の両辺の平均値をとれば，

$$E(Z_N) \leqq -(b-a)E(\beta_N) + E((X_N-a)^+).$$

したがって $E(Z_N) \geqq 0$ をいえば，定理の証明は終る．Z_N の定義により

$$E(X_{\sigma_{i+1}\wedge N}) \geqq E(X_{\tau_i\wedge N})$$

をいえばよい．X が $\boldsymbol{F}$ 劣マルチンゲールであり，$\tau_i \wedge N \leqq \sigma_{i+1}\wedge N \leqq N$ であるから，定理 5.13 により，$\tau_i\wedge N$, $\sigma_{i+1}\wedge N$ が $\boldsymbol{F}$ 停止時であることをいえば，上の不等式がでる．したがって，σ_i, τ_i が $\boldsymbol{F}$ 停止時であることをいえばよいことが定理 5.9 からわかる．

$$\{\sigma_1=n\} = \bigcap_{i<n}\{X_i>a\} \cap \{X_n \leqq a\} \in \mathscr{F}_n,$$

$$\{\tau_1=n\} = \bigcup_{j<n}\left(\{\sigma_1=j\} \cap \bigcap_{j<i<n}\{X_i<b\}\right) \cap \{X_n \geqq b\} \in \mathscr{F}_n,$$

$$\{\sigma_2=n\} = \bigcup_{j<n}\left(\{\tau_1=j\} \cap \bigcap_{j<i<n}\{X_i>a\}\right) \cap \{X_n \leqq a\} \in \mathscr{F}_n,$$

以下同様にして

$$\{\sigma_i=n\},\ \{\tau_i=n\} \in \mathscr{F}_n, \qquad i=1,2,\cdots.$$

これで σ_i, τ_i が $\boldsymbol{F}$ 停止時であることが証明された. ▌

定理 5.16 (マルチンゲールに関する収束定理) $X=\{X_n\}$ が $\boldsymbol{F}=\{\mathscr{F}_n\}$ に関してマルチンゲールとし, $\sup E(|X_n|)<\infty$ とすれば, $\{X_n\}$ は概収束し, 極限変数は可積分である. 劣マルチンゲール, 優マルチンゲールについても同様である.

証明 $\{X_n\}$ が優マルチンゲールならば, $\{-X_n\}$ が劣マルチンゲールであるから, 劣マルチンゲールのときに証明すればよい.

$$A = \{\liminf X_n < \limsup X_n\},$$
$$A_{a,b} = \{\liminf X_n < a < b < \limsup X_n\}$$

とすれば

$$A = \bigcup_{\substack{a<b \\ a,b\in \boldsymbol{Q}}} A_{a,b}.$$

$a<b$ を固定し, $[a,b]$ に対する, 時点 N までの X の上渡回数を $\beta_N=\beta_N(a,b)$ とすれば, 前定理により

$$E(\beta_N) \leqq \frac{E(|X_N|)+|a|}{b-a},$$

β_N は N とともに増大するから

$$E\left(\lim_{N\to\infty}\beta_N\right) = \lim_{N\to\infty}E(\beta_N) \leqq (b-a)^{-1}\left(\sup_N E(|X_N|)+|a|\right) < \infty.$$

したがって

$$P\left\{\lim_{N\to\infty}\beta_N=\infty\right\} = 0$$

が得られる. β_N, $\liminf X_n$, $\limsup X_n$ の定義により

$$\liminf X_n < a < b < \limsup X_n \Longrightarrow \lim \beta_N = \infty.$$

したがって

$$P\{A_{a,b}\} \leqq P\{\lim \beta_N=\infty\} = 0.$$

これから $P(A)=0$, すなわち

$$P\{\liminf X_n = \limsup X_n\} = 1$$

が得られる. これは $X_\infty=\lim X_n$ がほとんど確実に存在することを意味する. X_∞ は $\pm\infty$ となるかもしれない. しかし Fatou の定理により

$$E(|X_\infty|) = E(\lim |X_n|) \leqq \liminf E(|X_n|) < \infty$$

となるから, X_∞ が可積分となる. したがって当然 $X_\infty=\lim X_n$ はほとんど確実に有限である. これは $\{X_n\}$ の概収束を示している. ▌

上の定理で得られた極限変数 X_∞ は $\mathscr{F}_\infty=\bigvee_n \mathscr{F}_n$ に関して可測である. 実際すべての X_n は可測 $\mathscr{F}_n$ したがって可測 $\mathscr{F}_\infty$ であるから, $X_\infty=\lim X_n$ も可測 $\mathscr{F}_\infty$ である.

定理 5.17 $\{X_n\}$ が $\{\mathscr{F}_n\}$ に関してマルチンゲールであり, $\{X_n\}$ が一様可積分ならば, $\{X_n\}$ は概収束かつ1次平均収束し, その極限 X_∞ は $L^1(\mathscr{F}_\infty)\left(\mathscr{F}_\infty=\bigvee_n \mathscr{F}_n\right)$ に属し, かつ

$$X_n = E(X_\infty \mid \mathscr{F}_n)$$

がなりたつ. 劣(優)マルチンゲールのときには等号を $\leqq$ ($\geqq$) でおきかえたらよい.

証明 一様可積分ならば, 当然 $\sup E(|X_n|)<\infty$ であるから, 前定理が利用できて, $\{X_n\}$ は概収束し, その極限 X_∞ は $L^1(\mathscr{F}_\infty)$ に属する. したがって定理 3.24 により, $\{X_n\}$ は1次平均収束でもある. これから, $m>n$ のとき

$$E(|E(X_m \mid \mathscr{F}_n)-E(X_\infty \mid \mathscr{F}_n)|) \leqq E(|X_m-X_\infty|) \longrightarrow 0 \qquad (m\to\infty).$$

これは $E(X_m \mid \mathscr{F}_n)$, $m=n+1, n+2, \cdots$ が $E(X_\infty \mid \mathscr{F}_n)$ に1次平均収束することを意味する. したがって適当な列 $m_1, m_2, \cdots$ をとると

$$E(X_{m_p} \mid \mathscr{F}_n) \longrightarrow E(X_\infty \mid \mathscr{F}_n) \ \text{a.s.} \qquad (p\to\infty)$$

X が $\{\mathscr{F}_n\}$ 劣マルチンゲールならば, 左辺$\geqq X_n$ であるから,

$$X_n \leqq E(X_\infty \mid \mathscr{F}_n) \ \text{a.s.}$$

優マルチンゲール, マルチンゲールのばあいも同様である. ▌

いままで時間域 T が $\{0,1,2,\cdots\}$ のばあいを論じてきたが, $T=\{\cdots,-2,-1,0\}$ のばあいにもほぼ同様のことがいえる. $X=\{\cdots, X_{-2}, X_{-1}, X_0\}$ を確率過程と

し，$\boldsymbol{F}=\{\cdots,\mathcal{F}_{-2},\mathcal{F}_{-1},\mathcal{F}_0\}$ を増大列とする．

$$X_{-n}\in L^1(\mathcal{F}_{-n}),\qquad E(X_{-n}\mid\mathcal{F}_{-n-1})=X_{-n-1}\qquad(n=0,1,\cdots)$$

のとき，$\{X_{-n}\}$ は $\{\mathcal{F}_{-n}\}$ マルチンゲールという．劣(優)マルチンゲールに関しても同様である．

$\{X_{-n}\}$ が $\{\mathcal{F}_{-n}\}$ マルチンゲールならば，任意の $n\in\boldsymbol{N}$ に対して

$$Y_n{}^N=X_{(-N+n)\wedge 0},\qquad \mathcal{F}_n{}^N=\mathcal{F}_{(-N+n)\wedge 0},\qquad n=0,1,2,\cdots$$

とおくと，$Y^N=\{Y_0{}^N,Y_1{}^N,Y_2{}^N,\cdots\}$ は $\boldsymbol{F}^N=\{\mathcal{F}_0{}^N,\mathcal{F}_1{}^N,\mathcal{F}_2{}^N,\cdots\}$ に関して既述の意味のマルチンゲールである．(劣(優)マルチンゲールについても同様.) $\{X_{-n}\}$ が $\{\mathcal{F}_{-n}\}$ 劣マルチンゲールのときには，上の Y^N に上渡回数定理を適用して，

$$X_{-N},\ X_{-N+1},\ X_{-N+2},\ \cdots,\ X_0$$

が $[a,b]$ を上渡する回数 β_N に関してつぎの関係を得る．

$$E(\beta_N)\leqq\frac{E((X_0-a)^+)}{b-a}\leqq\frac{E(|X_0|)+|a|}{b-a}$$

したがって

$$E\Big(\lim_{N\to\infty}\beta_N\Big)<\infty$$

となる．これを用いて，定理5.16の証明と同様の論法により，

$$X_{-\infty}=\lim X_{-n}$$

がほとんど確実に存在することがわかる．$X_{-\infty}$ は $\pm\infty$ をとり得るが，実は $X_{-\infty}<\infty$ a.s. であることが示される．実際 $f(x)=x^+$ は増大凸関数でかつ $|f(x)|\leqq|x|$ であるから，$\{f(X_{-n})\}$ も $\{\mathcal{F}_{-n}\}$ 劣マルチンゲールである．したがって

$$E(f(X_{-n}))\leqq E(f(X_0)).$$

f は非負連続であるから，

$$E(f(X_{-\infty}))=E(\lim f(X_{-n}))\leqq\liminf E(f(X_{-n}))\leqq E(f(X_0))<\infty,$$

$$f(X_{-\infty})<\infty\ \text{a.s.}\quad すなわち\quad X_{-\infty}<\infty\ \text{a.s.}$$

$\{X_{-n}\}$ が $\{\mathcal{F}_{-n}\}$ 優マルチンゲールのときには，$\{-X_{-n}\}$ は $\{\mathcal{F}_{-n}\}$ 劣マルチンゲールとなるから，$-X_{-\infty}<\infty$ すなわち $X_{-\infty}>-\infty$ a.s. となる．

$\{X_{-n}\}$ が $\{\mathcal{F}_{-n}\}$ マルチンゲールのときには，つぎのようなよい結果が得られる．

定理 5.18 (マルチンゲールに関する逆向き収束定理) $X=\{X_{-n}\}$ が $\boldsymbol{F}=\{\mathcal{F}_{-n}\}$ に関してマルチンゲールであれば，$\{X_{-n}\}$ は一様可積分で，概収束かつ1次平均収束する．その極限 $X_{-\infty}$ は $\mathcal{F}_{-\infty}=\bigwedge\mathcal{F}_{-n}$ に関して可測で

$$X_{-\infty}=E(X_{-n}\mid\mathcal{F}_{-\infty}),\qquad n=0,1,2,\cdots.$$

証明 $\{X_{-n}\}$ も $\{-X_{-n}\}$ も劣マルチンゲールであるから，$\{X_{-n}\}$ は概収束し，その極限 $X_{-\infty}$ は

$$E(|X_{-\infty}|)=E((X_{-\infty})^+)+E((-X_{-\infty})^+)<\infty.$$

したがって，$X_{-\infty}$ はほとんど確実に有限である．

$$X_{-\infty}=\lim_{n\to\infty}X_{-n-k}$$

であるから，$X_{-\infty}$ は，任意の k に関して可測 $\mathcal{F}_{-k}$，したがって $\mathcal{F}_{-\infty}=\bigwedge\mathcal{F}_{-k}$ に関して可測である．1次平均収束をいうには，定理3.24により，$\{X_{-n}\}$ が一様可積分なことをいえばよい．$\{X_{-n}\}$ が $\{\mathcal{F}_{-n}\}$ マルチンゲールであるから，$\{|X_{-n}|\}$ は $\{\mathcal{F}_{-n}\}$ 劣マルチンゲールである．したがって

$$E\{|X_{-n}|:|X_{-n}|\geqq a\}\leqq E\{E(|X_0|\mid\mathcal{F}_{-n}):|X_{-n}|\geqq a\}$$
$$\leqq E(|X_0|,|X_{-n}|\geqq a),$$
$$P\{|X_{-n}|\geqq a\}\leqq E(|X_{-n}|)/a\leqq E(|X_0|)/a.$$

第2式から，$a\to\infty$ のとき，n に関して一様に $P\{|X_{-n}|\geqq a\}\to 0$．したがって n に関して一様に

$$E(|X_0|,|X_{-n}|\geqq a)\longrightarrow 0\qquad(a\to\infty).$$

これと第1の不等式により，n に関して一様に

$$E\{|X_{-n}|:|X_{-n}|\geqq a\}\longrightarrow 0\qquad(a\to\infty).$$

これは $\{X_{-n}\}$ の一様可積分性を示す．▮

例題 5.5 (i) 定理4.1 (Kolmogorov の不等式) を Doob の不等式 (定理5.15) より導け．

[ヒント] $X_1,X_2,\cdots,X_n$ が独立で $EX_i=0$ であるから $S_k=\sum_{i=1}^{k\wedge n}X_i$，$k=0,1,2,\cdots$ はマルチンゲールである．したがって $\{|S_k|^2\}$ は劣マルチンゲールである．

(ii) 定理4.3 (Kolmogorov の定理) をマルチンゲールの収束定理から導け．

[ヒント] $S_n=\sum_{i=1}^{n}(X_i-EX_i)$ がマルチンゲールであることと $E(|S_n|)^2\leqq E(S_n{}^2)\leqq\sum_{i=1}^{\infty}V(X_i)<\infty$ に注意せよ．

(iii) $E(|X|)<\infty$ ならば，$X_n=E(X\mid\mathcal{F}_n)$，$n=0,1,2,\cdots$ は一様可積分な $\{\mathcal{F}_n\}$

マルチンゲールであることを示せ．またこのとき

$$\lim_{n\to\infty} X_n = E(X \mid \mathcal{F}_\infty), \qquad \mathcal{F}_\infty = \bigvee \mathcal{F}_n$$

がなりたつことを示せ．

［ヒント］ $\{X_n\}$ の一様可積分性の証明は定理 5.18 のばあいと同様である．$\{X_n\}$ に定理 5.17 を適用すると $X_\infty = \lim X_n \in L^1(\mathcal{F}_\infty)$ の存在および $X_n = E(X_\infty \mid \mathcal{F}_n)$ までわかる．したがって，$E(X \mid \mathcal{F}_n) = E(X_\infty \mid \mathcal{F}_n)$，すなわち

$$E(X, A) = E(X_\infty, A), \qquad A \in \mathcal{F}_n.$$

$\mathcal{F}_\infty = \bigvee \mathcal{F}_n$ は乗法族 $\mathcal{A} = \bigcup \mathcal{F}_n$ で生成されるから，Dynkin 族定理により，上の等式が $A \in \mathcal{F}_\infty$ に対してなりたち，X_∞ が可測 $\mathcal{F}_\infty$ であることから，$X_\infty = E(X \mid \mathcal{F}_\infty)$ がでる．

(iv) $E(|X|) < \infty$ ならば，$X_{-n} = E(X \mid \mathcal{F}_{-n})$，$n = 0, 1, 2, \cdots$ は一様可積分な $\{\mathcal{F}_{-n}\}$ マルチンゲールであることを証明せよ．またこのとき

$$\lim_{n\to\infty} X_{-n} = E(X \mid \mathcal{F}_{-\infty}), \qquad \mathcal{F}_{-\infty} = \bigwedge \mathcal{F}_n$$

がなりたつことを示せ．

［ヒント］ 定理 5.18 を用いよ．

註 上の (iii), (iv) は，Doob がマルチンゲールの理論をつくる前から，P. Lévy の定理として知られていたものである．

(v) $\{X_n\}$ が $\{\mathcal{F}_n\}$ マルチンゲールで，$\{Y_n\}$ が $\{\mathcal{F}_n\}$ に適合して，しかも $Y_n \in L^\infty$ とすると

$$S_n = \sum_{k=0}^{n-1} Y_k (X_{k+1} - X_k) \qquad (S_0 = 0)$$

も $\{\mathcal{F}_n\}$ マルチンゲールであることを示せ．

(vi) $\{X_n\}$ が $\{\mathcal{F}_n\}$ 劣マルチンゲールで，$\sup_n E(|X_n|^p) < \infty$ $(p > 1)$ ならば，$\{X_n\}$ は概収束し，その極限 X_∞ に対して

$$X_\infty \in L^p(\mathcal{F}_\infty) \quad (\mathcal{F}_\infty = \bigvee \mathcal{F}_n), \qquad X_n \leqq E(X_\infty \mid \mathcal{F}_n)$$

がなりたつことを証明せよ．

［ヒント］ 例題 3.7 (i) により一様可積分性がでる．あとは定理 5.16, 5.17 を用いよ．

(vii) (定理 5.14 の変形) $\{X_n\}$ が $\{\mathcal{F}_n\}$ 劣マルチンゲールならば，$a > 0$ に対し

$$P\left\{\max_{k=0}^{N} X_k \geqq a\right\} \leqq \frac{1}{a} E(|X_N|),$$

$$P\left\{\min_{k=0}^{N} X_k \leqq -a\right\} \leqq \frac{1}{a} (E(|X_N|) - EX_0),$$

したがって

$$E\left(\max_{k=0}^{N} |X_k| \geqq a\right) \leqq \frac{1}{a} E(2|X_N| + |X_0|).$$

［ヒント］ 定理 5.14 の証明の Y_N と σ をとる．$\sigma \leqq N$ であるから，

$$\begin{aligned} P(Y_N \geqq a) &= P(X_\sigma \geqq a) \leqq \frac{1}{a} E\{X_\sigma, X_\sigma \geqq a\} \\ &\leqq \frac{1}{a} E(E(X_N \mid \mathcal{F}_\sigma), X_\sigma \geqq a) \\ &= \frac{1}{a} E(X_N, X_\sigma \geqq a) \leqq \frac{1}{a} E(|X_N|). \quad (\{X_\sigma \geqq a\} \in \mathcal{F}_\sigma,\ \text{定理 5.13}) \end{aligned}$$

同様に第 2 式の中の最小値を Z_N，

$$\tau = \begin{cases} \inf\{k \leqq N \mid X_k \leqq -a\} & (Z_N \leqq -a \text{ のとき}), \\ N & (Z_N > -a \text{ のとき}) \end{cases}$$

とすると，$0 \leqq \tau \leqq N$ により

$$\begin{aligned} E(X_0) &\leqq E(X_\tau) = E(X_\tau, X_\tau > -a) + E(X_\tau, X_\tau \leqq -a) \\ &\leqq E(X_N, X_\tau > -a) - aP\{X_\tau \leqq -a\} \\ &\leqq E(|X_N|) - aP\{Z_N \leqq -a\}. \end{aligned}$$

§5.6 連続時変数のマルチンゲール

T を区間とし，$X = \{X_t, t \in T\}$ を増大情報系 $\boldsymbol{F} = \{\mathcal{F}_t, t \in T\}$ に適合する確率過程とする．もし

$$E(|X_t|) < \infty, \qquad t \in T,$$
$$E(X_s \mid \mathcal{F}_t) = X_t, \qquad t, s \in T, \quad t < s$$

がなりたつとき，X を $\boldsymbol{F}$ マルチンゲールという．等号を $\geqq$ または $\leqq$ でおきかえて $\boldsymbol{F}$ 劣マルチンゲール，または $\boldsymbol{F}$ 優マルチンゲールを定義する．本節では $T = [0, \infty)$ のばあいを考えるが，これは本質的な制限ではない．

$\{X_t\}$ が $\{\mathcal{F}_t\}$ マルチンゲールとすると，$E(X_t)$ は一定である．劣マルチンゲールのときには，$E(X_t)$ は増大，優マルチンゲールのときには $E(X_t)$ は減少する．

f が区間 I の上で凸で

$$X_t(\omega) \in I \text{ a.s.}, \qquad E(|f(X_t)|) < \infty \qquad (t \in T)$$

とする．$\{X_t\}$ が $\{\mathcal{F}_t\}$ マルチンゲールならば，$\{f(X_t)\}$ は $\{\mathcal{F}_t\}$ 劣マルチンゲールである．$\{X_t\}$ が $\{\mathcal{F}_t\}$ 劣マルチンゲールで，f が増加関数ならば，$\{f(X_t)\}$ は $\{\mathcal{F}_t\}$ 劣マルチンゲールである．証明には，定理 5.12 と同様に，Jensen の不等式を用いる．

$\{X_t\}$ を $\{\mathcal{F}_t\}$ マルチンゲールとする．T の中に増大列 $\{t_n\}$ をとり

$$Y_n = X_{t_n}, \qquad \mathcal{G}_n = \mathcal{F}_{t_n}$$

とおくと，$\{Y_n\}$ は $\mathcal{G}_n$ マルチンゲールとなる．また $\{t_n\}$ が減少列のときには

$$Y_{-n} = X_{t_n}, \qquad \mathcal{G}_{-n} = \mathcal{F}_{t_n}$$

とおくと，$\{Y_{-n}\}$ は $\{\mathcal{G}_{-n}\}$ マルチンゲールとなる．このことは劣(優)マルチンゲールについても同様である．

定理 5.19 (**Doob の D 変形定理**) 任意の確率右連続な $\{\mathcal{F}_t\}$ 劣マルチンゲール $\{X_t\}$ に対し，これと同等な D 過程 $\{Y_t\}$ が存在する．

註 $\{X_t\}$ が確率右連続であるとは

$$\lim_{t \downarrow s} E(|X_t - X_s| \wedge 1) = 0, \qquad s \in T$$

となることである．もし D 過程 $\{Y_t'\}, \{Y_t''\}$ がともに $\{X_t\}$ と同等ならば，当然 $\{Y_t'\}, \{Y_t''\}$ は強同等である．ゆえに定理の $\{Y_t\}$ は強同等性を除いて一意に定まり，$\{X_t\}$ の **D 変形**という．$\{Y_t\}$ が $\{\mathcal{F}_t\}$ 劣マルチンゲールであることは定義から自明である．

定理 5.19 の証明 自然数 k に対して

$$Y_n^k = X_{n/2^k}, \qquad \mathcal{F}_n^k = \mathcal{F}_{n/2^k} \qquad (n=0, 1, 2, \cdots)$$

とおくと $Y^k = \{Y_0^k, Y_1^k, Y_2^k, \cdots\}$ は $\boldsymbol{F}^k = \{\mathcal{F}_0^k, \mathcal{F}_1^k, \mathcal{F}_2^k, \cdots\}$ に関して劣マルチンゲールである．N を任意の自然数とする．時点 2^kN までに Y^k が区間 $[a, b]$ を上渡する回数を β_N^k とすると，上渡回数定理により

$$E(\beta_N^k) \leqq \frac{E(|Y_{2^kN}^k|) + |a|}{b-a} = \frac{E(|X_N|) + |a|}{b-a}.$$

右辺は k に無関係で，これを c とおく．β_N^k は，$\{X_t\}$ を時点列

$$\Delta_N^k = \{0, 1/2^k, 2/2^k, \cdots, N\}$$

で観察したときに，区間 $[a, b]$ を上渡する回数である．N をきめたとき，Δ_N^k は k とともに増大するから，β_N^k も k とともに増大する．その極限を β_N^∞ とおくと

$$E(\beta_N^\infty) \leqq \lim E(\beta_N^k) \leqq c < \infty,$$

したがって

$$P\{\beta_N^\infty = \infty\} = 0.$$

さて

$$\Delta_N^k \uparrow \boldsymbol{Q}_2 \cap [0, N] \qquad (k \to \infty), \quad (\boldsymbol{Q}_2 = \{m \mid 2^k, m, k=1, 2, \cdots\})$$

に注意すると

$$(\exists t \in [0, N]) \quad \liminf_{r \uparrow t} X_r(\omega) < a < b < \limsup_{r \uparrow t} X_r(\omega)$$
$$\Longrightarrow \beta_N^\infty(\omega) = \infty \qquad (r \text{ は } \boldsymbol{Q}_2 \cap [0, N] \text{ の上を動く}).$$

さて $P\{\beta_N^\infty = \infty\} = 0$ であるから

$$P\left\{(\exists t \in [0, N]) \ \liminf_{r \uparrow t} X_r < a < b < \limsup_{r \uparrow t} X_r\right\} = 0.$$

a, b $(a<b)$ を有理数の範囲で動かして

$$P\left\{(\forall t \in [0, N]) \ \liminf_{r \uparrow t} X_r = \limsup_{r \uparrow t} X_r\right\} = 1$$

がいえることは，マルチンゲールの収束定理(定理 5.16)の証明と同様である．$\uparrow$ を $\downarrow$ でおきかえても，同じことがいえるから，Ω の中からある P 零集合 A_N をぬき去っておけば，すべての $t \in [0, N]$ に対して

$$\lim_{r \uparrow t} X_r, \quad \lim_{r \downarrow t} X_r \qquad (r \text{ は } \boldsymbol{Q}_2 \cap [0, N] \text{ の上を動く})$$

が確定する．これから $\omega \in \Omega - A \left(A = \bigcup_N A_N\right)$ に対しては

$$\lim_{r \uparrow t} X_r, \quad \lim_{r \downarrow t} X_r \qquad (r \in \boldsymbol{Q}_2)$$

が確定する．しかしこの極限は $\pm\infty$ であるかもしれない．この可能性を排除するために，例題 5.5 (vii) を用いて

$$P\left\{\max_{r \in \Delta_N^k} |X_r| \geqq a\right\} \leqq a^{-1} E(2|X_N| + |X_0|).$$

左辺の最大値は $k\to\infty$ のとき，増大して

$$\sup\{|X_r| \mid r\in \boldsymbol{Q}_2\cap[0,N]\}$$

に近づくから，

$$P\{\sup\{|X_r| \mid r\in \boldsymbol{Q}_2\cap[0,N]\}>a\} \leqq a^{-1}E(|X_N|+|X_0|),$$
$$P\{\sup\{|X_r| \mid r\in \boldsymbol{Q}_2\cap[0,N]\}=\infty\} = 0,$$
$$P\{(\exists N)\sup\{|X_r| \mid r\in \boldsymbol{Q}_2\cap[0,N]\}=\infty\} = 0.$$

この最後の括弧の中の事象（ω 集合）を B とすれば，$\omega\in\Omega-(A\cup B)$ に対しては

$$\lim_{r\uparrow t} X_r, \quad \lim_{r\downarrow t} X_r \qquad (r\in \boldsymbol{Q}_2)$$

は存在して，有限となる．

さて

$$Y_t(\omega)=\begin{cases}\lim\limits_{r\downarrow t} X_r(\omega), & \omega\in\Omega-(A\cup B),\\ 0, & \omega\in A\cup B\end{cases}$$

とおけば，任意の $\omega\in\Omega$ に対して $Y_t(\omega)$ は t の右連続関数である．また，$\omega\in\Omega-(A\cup B)$ に対し，$\lim\limits_{r\uparrow t} X_t(\omega)$ が有限確定することと，$\boldsymbol{Q}_2\cap[0,\infty)$ が $[0,\infty)$ で稠密なことから，すべての $\omega\in\Omega$ に対して，$\lim\limits_{s\uparrow t} Y_s(\omega)$ が有限確定することがわかる．かくして $\{Y_t\}$ が D 過程であることがわかった．

残る所は $\{Y_t\}$ と $\{X_t\}$ との同等性である．確率右連続性により任意に固定した t に対し適当な有理数列 $r_n\downarrow t$ をとると

$$X_t(\omega)=\lim_{n\to\infty} X_{r_n}(\omega) \ \text{a.s.}$$

右辺は $Y_t(\omega)$ の定義により，$Y_t(\omega)$ とほとんど確実に一致する．ゆえに

$$X_t(\omega)=Y_t(\omega) \ \text{a.s.} \qquad \blacksquare$$

上の定理を背景として，今後は D 過程であるような $\{\mathcal{F}_t\}$ 劣マルチンゲール（簡単に D 型の $\{\mathcal{F}_t\}$ 劣マルチンゲールという）を考えることが多い．

例題 5.6 （i） $\{X_t\}$ を $\{\mathcal{F}_t\}$ 劣マルチンゲールとし，

$$X_t\geqq 0 \ \text{a.s.}, \qquad t\in[0,\infty)$$

とする．このとき，任意の N に対し，$\{X_t, 0\leqq t\leqq N\}$ は一様可積分であることを示せ．

［ヒント］ $E(|X_t|, |X_t|\geqq a)=E(X_t, X_t\geqq a)\leqq E(X_N, X_t\geqq a),$
$P\{X_t\geqq a\}\leqq a^{-1}E(X_t)\leqq a^{-1}E(X_N).$

（ii） $\{X_t\}$ は (i) と同じとする．0 と N との間の離散値をとる $\{\mathcal{F}_t\}$ 停止時の全体を Σ_N とする．$\{X_\sigma, \sigma\in\Sigma_N\}$ は一様可積分であることを示せ．

［ヒント］ σ が $0=t_0<t_1<t_2<\cdots<t_n=N$ の中の値をとるとし

$$Y_k=X_{t_k}, \quad \mathcal{G}_k=\mathcal{F}_{t_k} \qquad (0\leqq k\leqq n),$$
$$Y_k=X_N, \quad \mathcal{G}_k=\mathcal{F}_N \qquad (k>n)$$

とすると，$\{Y_k\}$ は $\{\mathcal{G}_k\}$ 劣マルチンゲールである．$\sigma=t_k$ のとき $\theta=k$ とおくと，θ は $\{\mathcal{G}_k\}$ 停止時である．$\theta\leqq n$ であるから，定理 5.13 により

$$E(|Y_\theta|, |Y_\theta|\geqq a)=E(Y_\theta, Y_\theta\geqq a)\leqq E(Y_n, Y_\theta\geqq a),$$
$$E(|X_\sigma|, |X_\sigma|\geqq a)\leqq E(X_N, X_\sigma\geqq a),$$
$$P\{X_\sigma\geqq a\}\leqq a^{-1}E(X_\sigma)=a^{-1}E(Y_\theta)\leqq a^{-1}E(Y_n)=a^{-1}E(X_N).$$

（iii） $\{X_t\}$ は (i) と同じ条件をみたし，かつ D 型とする．任意の $\{\mathcal{F}_t\}$ 停止時 σ,τ $(0\leqq\sigma\leqq\tau\leqq N)$ に対し

$$E(X_\tau \mid \mathcal{F}_\sigma)\geqq X_\sigma$$

がなりたつことを示せ．また $0\leqq\sigma\leqq N$ をみたす任意の $\{\mathcal{F}_t\}$ 停止時に対する X_σ の全体は一様可積分であることを示せ．

［ヒント］ $\tau_n=n^{-1}([n\tau]+1), \quad \sigma_n=n^{-1}([n\tau]+1)$

は前問の Σ_{N+1} に入る．したがって X_{τ_n}, X_{σ_n}, $n=1,2,\cdots$ は一様可積分である．また $\tau_n\downarrow\tau$, $\sigma_n\downarrow\sigma$ であるから，

$$X_{\tau_n}\longrightarrow X_\tau, \quad X_{\sigma_n}\longrightarrow X_\sigma,$$

ゆえに

$$E(|X_{\tau_n}-X_\tau|)\longrightarrow 0, \quad E(|X_{\sigma_n}-X_\sigma|)\longrightarrow 0.$$

さて前問のように考えて

$$E(X_{\tau_n}, A)\geqq E(X_{\sigma_n}, A), \qquad A\in\mathcal{F}_\sigma\subset\mathcal{F}_{\sigma_n}.$$

したがって，$n\to\infty$ として

$$E(X_\tau, A)\geqq E(X_\sigma, A), \qquad A\in\mathcal{F}_\sigma.$$

（iv） $\{X_t\}$ が $\{\mathcal{F}_t\}$ マルチンゲールのときには，上の (i), (ii), (iii) がなりたつことを示せ．

［ヒント］ $\{|X_t|\}$ が $\{\mathcal{F}_t\}$ 劣マルチンゲールである．

§5.7　Gauss 系

実確率変数の系 $\mathcal{S}\subset L^0=L^0(\Omega, P)$ が Gauss 系であるとは，$\mathcal{S}$ の中の任意の有限個の1次結合が Gauss 分布に従うことである．$\mathcal{S}$ が Gauss 系ならば，任意の $X\in\mathcal{S}$ は Gauss 分布に従うべきであるから，当然 $X\in L^2=L^2(\Omega, P)$．したがって

$$\mathcal{S}\subset L^2$$

となる．

定理 5.20　$\mathcal{S}$ が Gauss 系であるための必要十分条件は，$\mathcal{S}$ の中の任意の有限個のものの結合変数が Gauss 分布に従うことである．

証明　$\mathcal{S}$ を Gauss 系とせよ．定義により

$$X=\sum_{k=1}^n t_kX_k,\qquad t_k\in \boldsymbol{R}^1,\quad X_k\in\mathcal{S}$$

が Gauss 分布に従う．したがって

$$E\left(\exp\left(iz\sum_{k=1}^n t_kX_k\right)\right)=\exp\left(imz-\frac{v}{2}z^2\right).$$

$$m=E(\sum t_kX_k)=\sum t_km_k,\qquad m_k=E(X_k),$$

$$v=V(\sum t_kX_k)=\sum_{k,l}t_kt_lv_{kl},\qquad v_{kl}=V(X_k, X_l)$$

$$\equiv E((X_k-m_k)(X_l-m_l)).$$

$z=1$ とおいて

$$E(\exp(i\sum t_kX_k))=\exp\left(i\sum m_kt_k-\frac{1}{2}\sum v_{kl}t_kt_l\right).$$

左辺は結合変数 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ の分布 $P^{\boldsymbol{X}}$ の特性関数，右辺は Gauss 分布 $N_{\boldsymbol{m},\boldsymbol{v}}$ の特性関数である．したがって $\boldsymbol{X}$ は Gauss 分布に従う．

逆に $\mathcal{S}$ の中の任意の有限個の元 $X_1, X_2, \cdots, X_n$ の結合変数 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ が Gauss 分布 $N_{\boldsymbol{m},\boldsymbol{v}}$ $(\boldsymbol{m}=(m_k),\ \boldsymbol{v}=(v_{kl}))$ に従うならば，$\{X_i\}$ の1次結合

$$X=\sum_{i=1}^n t_iX_i$$

の分布 μ に対して

$$(\mathcal{F}\mu)(z)=E\left(\exp\left(iz\sum_{k=1}^n t_kX_k\right)\right)=E\left(\exp\left(i\sum_{k=1}^n zt_kX_k\right)\right)$$

$$=(\mathcal{F}N_{\boldsymbol{m},\boldsymbol{v}})(zt_1, zt_2, \cdots, zt_n)$$

$$=\exp\left(iz\sum_k t_km_k-\frac{z^2}{2}\sum_{k,l}v_{kl}t_kt_l\right)$$

がなりたつ．したがって μ は Gauss 分布である．∎

さて $\mathcal{S}$ は L^2 の部分集合であり，L^2 は Hilbert 空間であるから，L^2 の中に $\mathcal{S}$ で張られるベクトル空間 $\mathcal{L}[\mathcal{S}]$ と，$\mathcal{L}[\mathcal{S}]$ の L^2 における閉包 $\overline{\mathcal{L}}[\mathcal{S}]$ が考えられる．

$\mathcal{S}$ が Gauss 系ならば，$\mathcal{S}$ の任意の部分系は Gauss 系である．また $\mathcal{L}[\mathcal{S}]$, $\overline{\mathcal{L}}[\mathcal{S}]$ も Gauss 系である．$\mathcal{L}[\mathcal{S}]$ の方は定義により明らかである．$\overline{\mathcal{L}}[\mathcal{S}]$ について証明する．

$$X_1, X_2, \cdots, X_n\in\overline{\mathcal{L}}[\mathcal{S}]$$

を任意にとって，その1次結合

$$Y=\sum t_kX_k$$

が Gauss 分布に従うことをいえばよい．

$$X_{1p}, X_{2p}, \cdots, X_{np}\in\mathcal{L}[\mathcal{S}],\qquad p=1, 2, \cdots$$

をとって

$$\|X_{kp}-X_k\|_2\longrightarrow 0\qquad (p\to\infty)$$

がなりたつようにできる．部分列をとることにより，同時に

$$X_{kp}\longrightarrow X_k\ \text{a.s.}\qquad (p\to\infty)$$

がなりたつようにできる．

$$Y_p=\sum_k t_kX_{kp}$$

とおくと，

$$\|Y_p-Y\|_2\longrightarrow 0,\qquad Y_p\longrightarrow Y\ \text{a.s.}$$

はじめの式から，

$$E(Y_p)=(Y_p, 1)\longrightarrow (Y, 1)=E(Y),$$

$$V(Y_p)=E(Y_p^2)-E(Y_p)^2=(Y_p, Y_p)-(Y_p, 1)^2$$

$$\longrightarrow E(Y^2)-E(Y)^2=V(Y).$$

'$Y_p\to Y$ a.s.' により

$$E(\exp(izY_p))\longrightarrow E(\exp(izY))\qquad (\text{有界収束定理}),$$

$$\text{左辺}=\exp\left(izE(Y_p)-\frac{z^2}{2}V(Y_p)\right)\longrightarrow\exp\left(izE(Y)-\frac{z^2}{2}V(Y)\right),$$

したがって

$$E(\exp(izY)) = \exp\left(izE(Y) - \frac{z^2}{2}V(Y)\right).$$

これは Y が Gauss 分布に従うことを意味する.

$\mathcal{S}$ を Gauss 系とし, $X_1, X_2 \in \mathcal{S}$ の共分散 $V(X_1, X_2)$ が 0 に等しいとすると

$$\begin{aligned} &E(\exp(iz_1X_1 + iz_2X_2)) \\ &= \exp\left\{iz_1m_1 + iz_2m_2 - \frac{1}{2}z_1^2v_1 - \frac{1}{2}z_2^2v_2\right\} \qquad (m_k = EX_k,\ \ v_k = V(X_k)) \\ &= E(\exp(iz_1X_1))E(\exp(iz_2X_2)). \end{aligned}$$

これから (X_1, X_2) の分布が X_1 の分布と X_2 の分布との直積であることがわかる. すなわち Gauss 系 $\mathcal{S}$ の中では2変数の共分散0は独立を意味する.

同じ方法で Gauss 系 $\mathcal{S}$ の中に $X_1, X_2, \cdots$ があり, 任意の X_i, X_j $(i \neq j)$ の共分散が0であると, $X_1, X_2, \cdots$ が独立となる.

この事実をさらに一般化して, 次の定理が得られる.

定理 5.21 $\mathcal{S}$ が Gauss 系で, $\mathcal{S}_1, \mathcal{S}_2, \cdots \subset \mathcal{S}$ とし, 任意の $i \neq j$ に対し

$$V(X, Y) = 0, \qquad X \in \mathcal{S}_i,\ \ Y \in \mathcal{S}_j$$

とする. このとき $\mathcal{S}_1, \mathcal{S}_2, \cdots$ は独立である. (これは情報系 $\mathcal{F}[\mathcal{S}_1], \mathcal{F}[\mathcal{S}_2], \cdots$ が独立であるといってもよい.) ——

証明の考え方は上述の特別の場合と同じであるから, 省略する.

上の定理は Gauss 系の著しい特色をあらわしている. 一般の確率変数列 $X_1, X_2, \cdots$ のときには, $\{X_1, X_2, \cdots\}$ が独立ということと, その中の任意の2個が独立である(このときに**相互に独立**という)ことは異なっていて, 前者は後者より強い条件である. しかし $X_1, X_2, \cdots$ がある Gauss 系の中に入っているときには, この2条件は同値であり, 実はともに '$V(X_i, X_j) = 0$ $(i \neq j)$' と同値である.

$\mathcal{S}$ が Gauss 系で, しかも $E(X) = 0$, $X \in \mathcal{S}$ のときには $\mathcal{S}$ を**平均 0 の Gauss 系**という. これは $L_0^2 \equiv L^2 \ominus \{1\}$ (1 の L^2 における直交余空間) に含まれる Gauss 系といってもよい. $\mathcal{S}$ が一般の Gauss 系のときには

$$\mathcal{S}_0 = \{X - E(X), X \in \mathcal{S}\}$$

は平均0の Gauss 系である. したがって平均0の Gauss 系の性質を調べると一般の Gauss 系の性質が容易に導かれる. さて L_0^2 の中では

$$V(X, Y) = E(X, Y) = (X, Y)$$

となるから, 共分散 $V(X, Y)$ が0ということは, X, Y が直交することになり, Hilbert 空間の言葉が用いやすい. 例えば $X_1, X_2, \cdots$ が平均0の Gauss 系 $\mathcal{S}$ の中にあるときには, これが相互に直交することと, $\{X_1, X_2, \cdots\}$ が独立なこととは同等である. $\mathcal{S}_0$ が平均0の Gauss 系ならば, $\mathcal{L}[\mathcal{S}], \bar{\mathcal{L}}[\mathcal{S}]$ も平均0の Gauss 系である.

任意の確率変数系 $\mathcal{S}$ に対し, $\mathcal{S}$ で生成される情報 $\mathcal{F}[\mathcal{S}]$ に関して可測な $X \in L^2$ の全体を $L^2[\mathcal{S}]$ であらわすこと, また $E(X|\mathcal{F}[\mathcal{S}])$ を $E(X|\mathcal{S})$ とかくことは前にのべた. 定理 3.32 によれば

$$E(X \mid \mathcal{S}) = p_{L^2[\mathcal{S}]/L^2}X.$$

ここで $p_{M_1/M}$ は Hilbert 空間 M からその閉部分空間 M_1 への射影をあらわす. Gauss 系の中ではこの関係が次のように簡単になる.

定理 5.22 $\mathcal{S}$ を平均0の Gauss 系, $\mathcal{T}$ をその部分系とすると,

$$E(X \mid \mathcal{T}) = p_{\bar{\mathcal{L}}[\mathcal{T}]/\bar{\mathcal{L}}[\mathcal{S}]}X, \qquad X \in \mathcal{S}.$$

証明 右辺を Y とおくと, $Y \in \bar{\mathcal{L}}[\mathcal{T}]$ であるから, 平均0の Gauss 分布に従う. 射影の定義により

$$X = Y + Z, \qquad Z \equiv Y - X \in \bar{\mathcal{L}}[\mathcal{S}] \ominus \bar{\mathcal{L}}[\mathcal{T}].$$

$\ominus$ は直交余空間をあらわす. Y は $\bar{\mathcal{L}}[\mathcal{T}]$ に属するから, Y と Z とは直交し, ともに $\bar{\mathcal{L}}[\mathcal{S}]$ に属する. $\bar{\mathcal{L}}[\mathcal{S}]$ は平均0の Gauss 系であるから, Y と Z とは独立である. Z は $\bar{\mathcal{L}}[\mathcal{T}]$ $(\subset \bar{\mathcal{L}}[\mathcal{S}])$ とも直交するから, Z は $\bar{\mathcal{L}}[\mathcal{T}]$ と独立である. したがって $\mathcal{T}$ とも独立である. これから Z が $\mathcal{F}[\mathcal{T}]$ と独立なこともわかる. したがって

$$E(Z \mid \mathcal{T}) = E(Z \mid \mathcal{F}[\mathcal{T}]) = E(Z) = 0.$$

(最後からふたつ目の等式は $A \in \mathcal{F}[\mathcal{T}]$ に対し

$$\begin{aligned} E(E(Z \mid \mathcal{F}[\mathcal{T}]), A) &= E(Z, A) = E(Z, 1_A) = E(Z)E(1_A) \\ &= E(E(Z), A) \end{aligned}$$

からすぐにわかる.)

上に得た等式から

$$E(X \mid \mathcal{T}) = E(Y \mid \mathcal{T}) + E(Z \mid \mathcal{T}) = E(Y \mid \mathcal{T}).$$

Y は $\mathcal{F}[\mathcal{T}]$ に関して可測であるから, 右辺は Y に等しい. ▌

この定理の効用を示すために，簡単な例をあげよう．

(X, Y) が2次元の Gauss 分布 $N_{0,v}$ に従うとし，

$$v = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} \qquad (0<r<1)$$

とする．このとき $E(Y|X)=rX$ を証明しよう．

もし定理3.32を用いるとすれば

$$E(Y|X) = p_{L^2[X]/L^2}Y$$

となる．さて

$$L^2[X] = \{f(X)\,|\,E(|f(X)|^2)<\infty\} = \left\{f(X)\,\middle|\,\int_{R^1}|f(x)|^2P^X(dx)<\infty\right\}$$
$$= \{f(X)\,|\,f\in L^2(\boldsymbol{R}^1, P^X)\}$$

であるから，$L^2[X]$ を考えるために $L^2(\boldsymbol{R}^1, P^X)$ に属する関数 f をすべて考える必要がおこる．さらに $L^2[X]$ への射影は $L^2=L^2(\Omega, P)$ の中で考えなければならない．

しかし $\{X, Y\}$ が平均0の Gauss 系であることに着目して，上の定理5.22を用いると，

$$E(Y|X) = p_{\bar{\mathscr{Z}}[X]/\bar{\mathscr{Z}}[X,Y]}Y$$

となり，

$$\bar{\mathscr{Z}}[X] = \{aX\,|\,a\in\boldsymbol{R}^1\},$$
$$\bar{\mathscr{Z}}[X, Y] = \{bX+cY\,|\,b, c\in\boldsymbol{R}^1\}$$

となる．したがって $E(Y|X)$ は aX の形をしていて

$$Y-aX\perp X$$

となっていることがわかる．

$$0 = (Y-aX, X) = (Y, X)-a(X, X) = r-a$$

となり $a=r$ となる．

さて確率論からはなれて，一般に対称正定行列

$$k = (k_{\alpha\beta}, (\alpha, \beta)\in A\times A)\in\boldsymbol{R}^{A\times A}$$

が与えられたとき，k を再生核とする Hilbert 空間 $H(k)$ を下のように定義する．$k_{\alpha\beta}$ で α をとめて β の関数とみたとき，$k_\alpha=k_\alpha(\beta)$ であらわす．k_α は関数空間 $\boldsymbol{R}^A$ の点であるから，$\{k_\alpha, \alpha\in A\}$ で張られる $\boldsymbol{R}^A$ の部分ベクトル空間が定まり，

これを $\mathscr{L}(k)$ であらわす．任意の $f\in\mathscr{L}(k)$ は

$$f = \sum_\alpha \xi_\alpha k_\alpha \qquad (\text{有限個の}\ \alpha\ \text{を除いて}\ \xi_\alpha=0)$$

の形にあらわされる．この表現は一意とは限らない．さて $f=\sum\xi_\alpha k_\alpha$，$g=\sum\eta_\alpha k_\alpha$ の内積 (f, g) を

$$(f, g) = \sum_{\alpha,\beta}\xi_\alpha\eta_\beta k_{\alpha\beta}$$

で定義する．右辺は一見無限和に見えるが，有限個の $(\alpha, \beta)\in A\times A$ を除いて $\xi_\alpha\eta_\beta=0$ となるから，実質的には有限和で，収束問題は考慮する必要はない．問題は (f, g) が f, g の表現に無関係に確定するかどうかである．実際 f, g が別の表現

$$f = \sum\xi_\alpha' k_\alpha, \qquad g = \sum\eta_\alpha' k_\alpha$$

をもつとすれば，上の内積の定義は

$$(f, g) = \sum_{\alpha,\beta}\xi_\alpha'\eta_\beta' k_{\alpha\beta}$$

となるから，これが前の定義と一致することをいえば，(f, g) の定義の意味が確定する．

$$\sum_{\alpha,\beta}\xi_\alpha\eta_\beta k_{\alpha\beta} = \sum_\beta\eta_\beta\sum_\alpha\xi_\alpha k_\alpha(\beta) = \sum_\beta\eta_\beta f(\beta)$$

であるから，

$$\sum_{\alpha,\beta}\xi_\alpha\eta_\beta k_{\alpha\beta} = \sum_{\alpha,\beta}\xi_\alpha'\eta_\beta k_{\alpha\beta} \quad \left(=\sum_\beta\eta_\beta f(\beta)\right)$$

となることは明らかである．同様に

$$\sum_{\alpha,\beta}\xi_\alpha'\eta_\beta k_{\alpha\beta} = \sum_{\alpha,\beta}\xi_\alpha'\eta_\beta k_{\beta\alpha} = \sum_\alpha\xi_\alpha' g(\alpha)$$

に注意して

$$\sum_{\alpha,\beta}\xi_\alpha'\eta_\beta k_{\alpha\beta} = \sum_{\alpha,\beta}\xi_\alpha'\eta_\beta' k_{\alpha\beta} \quad \left(=\sum_\alpha\xi_\alpha' g(\alpha)\right)$$

が得られ，結局

$$\sum\xi_\alpha\eta_\beta k_{\alpha\beta} = \sum\xi_\alpha'\eta_\beta' k_{\alpha\beta}$$

となり，(f, g) の意味が確定する．とくに上の計算により

$$(k_\alpha, k_\beta) = (k_\beta, k_\alpha) = k_{\alpha\beta} = k_\alpha(\beta) = k_\beta(\alpha),$$
$$(f, k_\beta) = (k_\beta, f) = f(\beta)$$

が得られる．

(f,g) の双1次性は明らか．$k_{\alpha\beta}$ の対称性から (f,g) の対称性がでる．また $(f,f)\geqq 0$ は $k=(k_{\alpha\beta})$ の正定性よりすぐにでる．これから

$$(f,g)^2\leqq(f,f)(g,g)$$

がでる．したがって

$$(f,f)=0 \Longrightarrow (f,g)=0,\quad g\in\mathcal{L}(k)$$
$$\Longrightarrow f(\alpha)=(f,k_\alpha)=0,\quad \alpha\in A.$$

これで (f,g) が内積のすべての性質をもち，$\mathcal{L}(k)$ にこの内積をそえて内積空間が得られる．

内積空間 $\mathcal{L}(k)$ の完備化を $H(k)$ とすれば，$H(k)$ は Hilbert 空間となる．$\mathcal{L}(k)$ は $\boldsymbol{R}^A$ の部分空間であるが，$H(k)$ は $\mathcal{L}(k)$ から抽象論法で得られたものであるから，$\boldsymbol{R}^A$ の部分空間というわけではない．しかし写像

$$\varphi:\ H(k)\longrightarrow \boldsymbol{R}^A,\quad f\longmapsto((f,k_\alpha),\alpha\in A)$$

を考えると，線型単射となっている．線型性は明らかである．$\{k_\alpha,\alpha\in A\}$ が $\mathcal{L}(k)$ を張り，$\mathcal{L}(k)$ が $H(k)$ の中で稠密であるから

$$\varphi f=0\Longrightarrow(\forall\alpha)\ (f,k_\alpha)=0\Longrightarrow f=0$$

となり，φ の単射性が得られる．

$f\in\mathcal{L}(k)$ に対しては，

$$f(\alpha)=(f,k_\alpha)$$

がなりたっているから，$\varphi:H(k)\to\boldsymbol{R}^A$ は $\mathcal{L}(k)$ $(\subset H(k))$ の上では恒等写像である．したがって φ によって $H(k)$ を $\boldsymbol{R}^A$ の中に埋めこむと

$$\mathcal{L}(k)\subset H(k)\subset\boldsymbol{R}^A$$

となり，当然

$$f(\alpha)(=(\varphi f)(\alpha))=(f,k_\alpha),\qquad \alpha\in A,\quad f\in H(k)$$

が得られる．$H(k)$ は Hilbert 空間であるが，線型空間としては $\boldsymbol{R}^A$ の部分空間であることは φ の線型性からすぐにわかる．$\mathcal{L}(k)$ は $H(k)$ の中で $(H(k)$ のノルムに関して$)$ 稠密である．さらに $\{f_n\}\subset H(k)$ が $f\in H(k)$ にノルムに関して収束すれば，$\{f_n\}$ は f に各点収束する．実際，$n\to\infty$ のとき

$$|f_n(\alpha)-f(\alpha)|=|(f_n,k_\alpha)-(f,k_\alpha)|\leqq\|f_n-f\|\|k_\alpha\|\longrightarrow 0.$$

しかし $\{f_n\}\subset H(k)$ が $f\in\boldsymbol{R}^A$ に各点収束しても，f は $H(k)$ に入るとは限らないし，またかりに f が $H(k)$ に入っても $\{f_n\}$ が f にノルム収束するとは限らない．

上のように定義された Hilbert 空間 $H(k)$ を k を**再生核**とする Hilbert 空間という．再生核の名称は，$H(k)$ の性質

$$(f,k_\alpha)=f(\alpha)$$

に由来する．

話を Gauss 系に戻そう．

$$\mathcal{S}=\{X_\alpha,\alpha\in A\}$$

を Gauss 系とするとき，

$$m=m(\mathcal{S})=(m_\alpha\equiv E(X_\alpha),\alpha\in A),$$
$$v=v(\mathcal{S})=(v_{\alpha\beta}\equiv V(X_\alpha,X_\beta),(\alpha,\beta)\in A\times A)$$

をそれぞれ $\mathcal{S}$ の**平均値ベクトル**，**分散行列**という．

分散行列 $v=v(\mathcal{S})$ が対称であることは明らかであるが，

$$\sum_{i,j}\xi_i\xi_j V(X_{\alpha_i},X_{\alpha_j})=E\left(\left(\sum_i\xi_i(X_{\alpha_i}-m_{\alpha_i})\right)^2\right)\geqq 0$$

により，v は正定符号である．したがって v を再生核とする Hilbert 空間 $H(v)$ が定まる．

上の $\mathcal{S}$ に対し

$$\mathcal{S}_0=\{X_\alpha-m_\alpha,\alpha\in A\}$$

とおくと，$\mathcal{S}_0$ は平均0の Gauss 系であり，Hilbert 空間 $L^2=L^2(\Omega,P)$ の部分集合である．$\mathcal{S}_0$ から $H(v)$ の中へ写像

$$\Phi:\ X_\alpha-m_\alpha\longrightarrow v_\alpha\qquad(v_\alpha(\beta)=v_{\alpha\beta})$$

は内積をかえない．すなわち

$$(X_\alpha-m_\alpha,X_\beta-m_\beta)=v_{\alpha\beta}=(v_\alpha,v_\beta).$$

したがって Φ は $\mathcal{S}_0$ で張られる L^2 の閉部分空間 $\overline{\mathcal{L}}[\mathcal{S}_0]$ から $\{v_\alpha,\alpha\in A\}$ で張られる $H(v)$ の閉部分空間(実は $H(v)$ 自身)への (Hilbert 空間としての)同型写像に拡張される．

さて (Ω,P) は完備可分な確率空間であるから，$L^2=L^2(\Omega,P)$ は可分な Hilbert 空間，したがって $\overline{\mathcal{L}}[\mathcal{S}_0]$ $(\subset L^2)$ も可分である．上の同型性により，$H(v)$ もまた可分でなければならない．

$H(v)$ の正規直交基を $e_1,e_2,\cdots$ とすると，$Z_n=\Phi^{-1}e_n$，$n=1,2,\cdots$ は $\overline{\mathcal{L}}[\mathcal{S}_0]$ の

正規直交基となる．$\mathcal{S}_0$ が平均0の Gauss 系であるから $\overline{\mathcal{L}}[\mathcal{S}_0]$ も平均0の Gauss 系である．したがって $\{Z_n\}\subset\overline{\mathcal{L}}[\mathcal{S}_0]$ の直交性から，その独立性がでる．$v_\alpha\in H(v)$ を $\{e_n\}$ で展開して

$$v_\alpha=\sum_n(v_\alpha, e_n)\cdot e_n$$

とすると，両辺に Φ^{-1} を施して

$$X_\alpha-m_\alpha=\sum_n(v_\alpha, e_n)Z_n$$

が得られる．この無限和はノルム収束(すなわち2次平均収束)しているのであるが，$\{Z_n\}$ の独立性を考慮すると，P. Lévy の3収束同等定理により，概収束していることにもなる．

$\overline{\mathcal{L}}[\mathcal{S}_0]$ が平均0の Gauss 型であるから，$V(Z_n)=(Z_n, Z_n)=1$ に注意すると Z_n の分布は $N_{0,1}$ であることがわかる．したがって確率ベクトル $(Z_1, Z_2, \cdots)$ の確率法則は

$$N_{0,1}{}^\infty=N_{0,1}\times N_{0,1}\times\cdots$$

に等しい．これだけの考察を念頭におくと，つぎの定理が得られる．

定理 5.23 (**Gauss 系の存在定理**) (i) $\mathcal{S}=\{X_\alpha, \alpha\in A\}$ を Gauss 系とし，$m=(m_\alpha)$，$v=(v_{\alpha\beta})$ をそれぞれ $\mathcal{S}$ の平均値ベクトル $m(\mathcal{S})$，分散行列 $v(\mathcal{S})$ とすると，v は対称正定型で，v を再生核とする Hilbert 空間 $H(v)$ は可分である．

(ii) $m=(m_\alpha, \alpha\in A)$ を任意のベクトル，$v=(v_{\alpha\beta}, (\alpha, \beta)\in A\times A)$ を任意の対称正定型の行列とし，v を再生核とする Hilbert 空間 $H(v)$ が可分であるとすれば，Gauss 系 $\mathcal{S}=\{X_\alpha, \alpha\in A\}$ が適当な確率空間 (Ω, P) の上に存在して，$m(\mathcal{S})=m$，$v(\mathcal{S})=v$ となる．しかもこのような $\mathcal{S}$ は法則同等を除いて一意である．すなわち別に (Ω', P') の上に Gauss 系 $\mathcal{S}'=\{X_\alpha', \alpha\in A\}$ があれば，任意の $\alpha_1, \alpha_2, \cdots, \alpha_n$ に対して $(X_{\alpha_1}', X_{\alpha_2}', \cdots, X_{\alpha_n}')$ と $(X_{\alpha_1}, X_{\alpha_2}, \cdots, X_{\alpha_n})$ とは同じ確率法則をもつ．

証明 (i)は上述の説明に含まれている．(ii)の証明をしよう．$H(v)$ の正規直交基 $\{e_n\}$ をとると

$$\sum_n(v_\alpha, e_n)^2=\|v_\alpha\|^2<\infty.$$

また

$$\Omega=\boldsymbol{R}^\infty, \qquad P=N_{0,1}{}^\infty,$$
$$Z_n(\omega)=\omega_n, \qquad \omega=(\omega_1, \omega_2, \cdots)\in\Omega$$

とおくと，$\{Z_n\}$ は (Ω, P) の上の独立確率変数列で，各変数は $N_{0,1}$ に従っている．したがって $\{Z_n\}$ は平均0の Gauss 系である．また $\{Z_n\}$ は $L^2=L^2(\Omega, P)$ の中の正規直交列である．

$$X_\alpha=m_\alpha+\sum_n(v_\alpha, e_n)Z_n$$

とおくと，$\sum_n(v_\alpha, e_n)^2<\infty$ により，上の級数は $L^2=L^2(\Omega, P)$ の中でノルム収束する．(したがって概収束もする.)

$$(X_\alpha-m_\alpha, X_\beta-m_\beta)=\sum_n(v_\alpha, e_n)(v_\beta, e_n)=(v_\alpha, v_\beta)=v_{\alpha\beta}.$$

また $EZ_n=0$ により

$$EX_\alpha=m_\alpha.$$

したがって

$$V(X_\alpha, X_\beta)=(X_\alpha-m_\alpha, X_\beta-m_\beta)=v_{\alpha\beta}$$

となる．$\{Z_n\}$ が Gauss 系であり，$X_\alpha-m_\alpha$ はすべて $\overline{\mathcal{L}}(\{Z_n\})$ に属するから，$\mathcal{S}=\{X_\alpha, \alpha\in A\}$ も Gauss 系である．上に示した所により，$m(\mathcal{S})=m$, $v(\mathcal{S})=v$ となる．これで存在が示された．一意性を示すには，定理5.20により，両変数の確率法則がともに Gauss 分布

$$N_{a,b} \qquad (a=(m_{\alpha_i}, i=1, 2, \cdots, n),\ \ b=(v_{\alpha_i\alpha_j}, i, j=1, 2, \cdots, n))$$

となることを注意すればよい．∎

例題 5.7 (i) 下の $(v_{\alpha\beta}, (\alpha, \beta)\in A\times A)$ は対称正定であることを示せ．

$$A=\boldsymbol{R}^1, \qquad v_{\alpha\beta}=e^{-|\alpha-\beta|},$$
$$A=\boldsymbol{R}^1, \qquad v_{\alpha\beta}=e^{-|\alpha-\beta|^2},$$
$$A=\boldsymbol{R}^1, \qquad v_{\alpha\beta}=(\mathcal{F}\mu)(\alpha-\beta) \qquad (\mu \text{ は } \boldsymbol{R}^1 \text{ 上の対称分布}).$$

[ヒント] 始めの2問は第3問の特別の場合である．

(ii) (E, μ) を測度空間とし，$\mathcal{A}$ を μ 測度有限の集合の全体とする．このとき

$$v_{\alpha\beta}=\mu(\alpha\cap\beta), \qquad \alpha, \beta\in\mathcal{A}$$

は対称正定であることを示せ．

[ヒント] $$v_{\alpha\beta}=\int_E 1_\alpha(x)1_\beta(x)\mu(dx).$$

(iii) 前問において $(v_{\alpha\beta})$ を再生核とする Hilbert 空間 $H(v)$ は Hilbert 空間 $L^2(E, \mu)$ と同型であることを示せ．

［ヒント］ $v_\alpha(\beta)=v_{\alpha\beta}$ とすると $\{v_\alpha, \alpha\in\mathcal{A}\}$ は $H(v)$ を張る．また $\{1_\alpha, \alpha\in\mathcal{A}\}$ も $L^2(E,\mu)$ を張る．しかも

$$(v_\alpha, v_\beta) = v_{\alpha\beta} = (1_\alpha, 1_\beta).$$

これから $H(v)$ と $L^2(E,\mu)$ の同型性がでる．

(iv) $X_1, X_2, \cdots, X_n$ が独立で，同じ Gauss 分布 $N_{m,v}$ に従うとするとき，

$$\bar{X} = \frac{1}{n}(X_1+X_2+\cdots+X_n),$$

$$S = \left[\frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2\right]^{1/2}$$

は独立であることを示せ．また $(\bar{X}, S)$ の分布を求めよ．

［ヒント］ X_i のかわりに $(X_i-m)/\sqrt{v}$ を考えることにより，$m=0$，$v=1$ のばあいに帰着される．独立性をいうには $(X_i, X_j)=\delta_{ij}$ に注意して

$$(\bar{X}, X_i-\bar{X}) = 0, \qquad i=1,2,\cdots,n$$

を示し，$(\bar{X}, X_1-\bar{X}, X_2-\bar{X}, \cdots, X_n-\bar{X})$ が Gauss 系であることに注意して，$\bar{X}$ と $(X_1-\bar{X}, X_2-\bar{X}, \cdots, X_n-\bar{X})$ の独立なことがいえる．S は後者の関数(娘)であるから，$\bar{X}$ と S とも独立である．$\boldsymbol{u}_1=(u_1{}^i=1/\sqrt{n}, i=1,2,\cdots,n)$ は $\boldsymbol{R}^n$ の中の単位ベクトルである．これに直交し，互いに直交する単位ベクトル $\boldsymbol{u}_\alpha=(u_\alpha{}^i, i=1, 2, \cdots, n)$ を任意につけ加えて，$\boldsymbol{R}^n$ の中の正規直交基 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ を定める．

$$Y_\alpha = \sum_{i=1}^n u_\alpha{}^i X_i$$

とおく．明らかに $Y_1=\sqrt{n}\,\bar{X}$ である．また

$$\sum_{\alpha=1}^n Y_\alpha{}^2 = \sum_{i,j}\left(\sum_\alpha u_\alpha{}^i u_\alpha{}^j\right)X_iX_j = \sum_i X_i{}^2 = n(\bar{X}^2+S^2),$$

ゆえに $\sum_{\alpha=2}^n Y_\alpha{}^2=nS^2$.

$$(Y_\alpha, Y_\beta) = \sum_{i,j} u_\alpha{}^i u_\beta{}^j (X_i, X_j) = \sum_i u_\alpha{}^i u_\beta{}^i = \delta_{\alpha\beta}$$

である．$\{X_n\}$ は平均0の Gauss 系で $\{Y_\alpha\}\subset\mathcal{L}\{X_n\}$ であるから，上の式から，$Y_1, Y_2, \cdots, Y_n$ は独立であること，各変数が Gauss 分布 $N_{0,1}$ に従うことがわかる．このことからも $\bar{X}=Y_1/\sqrt{n}$ と $S=\left(n^{-1}\sum_{\alpha=2}^n Y_\alpha{}^2\right)^{1/2}$ の独立性がわかる．したがって $(\bar{X}, S)$ の分布を求めるには，$\bar{X}$ の分布と S の分布との直積をとればよい．

$\bar{X}$ の分布は明らかに $N_{0,1/n^2}$ である．S の分布を求めるには，$A\in\mathcal{B}^1$ に対し

$$P\{S\in A\} = E(1_A(S))$$

$$= \int\cdots\int (2\pi)^{-(n-1)/2}\exp\left\{-\frac{1}{2}\sum_2^n y_\alpha{}^2\right\}1_A\left(n^{-1/2}\left(\sum_2^n y_\alpha{}^2\right)^{1/2}\right)dy_2dy_3\cdots dy_n$$

$$= c\int_0^\infty e^{-r^2/2}1_A(n^{-1/2}r)r^{n-2}dr \qquad (c\ \text{は定数})$$

$$= c'\int_0^\infty e^{-ns^2/2}1_A(s)s^{n-2}ds \qquad (c'\ \text{は定数}).$$

ゆえに S の分布は密度 $f(s)=c'e^{-ns^2/2}s^{n-2}$ をもつ．c' を求めるには，f の $[0,\infty)$ の上の積分が1に等しいことを用いよ．

§5.8 Wiener 過程 (Brown 運動)

Wiener 過程は1920年代に N. Wiener が Brown 運動の数学的模型として導入したもので，Wiener の Brown 運動ともよばれる．Wiener 過程の導入方法はいくつもあるが，たがいに同値である．ここでは Gauss 系として導入することにする．

確率過程 $\{X_t, t\in T\}$ が Gauss 系であるとき，**Gauss 過程**とよぶ．**平均0の Gauss 過程**も同様に定義する．

$B=\{B_t, t\in[0,\infty)\}$ が Wiener 過程であるとは，つぎの3条件がみたされることである．

(B.1) B は Gauss 過程である．

(B.2) $E(B_t)=0, \qquad V(B_t, B_s)=t\wedge s.$

(B.3) B は連続過程である．

始めの2条件だけを仮定したときには，**広義の Wiener 過程**という．

広義の Wiener 過程の存在と (法則同等を除いての) 一意性をいうには，

$$v = (t\wedge s, (t,s)\in[0,\infty)\times[0,\infty))$$

が対称正定で，$H(v)$ が可分であることをいえばよい (Gauss 系の存在定理)．対称性は明らか．正定性は

$$t\wedge s = \int_0^\infty 1_{[0,t]}(u)1_{[0,s]}(u)\,du$$

から得られる．また例題5.7 (iii) の論法を用いて，$H(v)$ が $L^2[0,\infty)$ と同型であ

ることがわかるから，$H(v)$ は可分である．

Wiener 過程の存在を証明するには，上に得られた広義の Wiener 過程を適当に変形すればよいのであるが，そのために準備としてつぎの定理を証明する．

定理 5.24 (Kolmogorov の連続変形定理)　$X=\{X_t, 0\leqq t\leqq a\}$ $(a<\infty)$ が

$$E(|X_t-X_s|^\alpha)\leqq\beta|t-s|^{1+\gamma},\quad 0\leqq s,t\leqq a\qquad (\alpha,\beta,\gamma \text{ は正の定数})$$

をみたすとき，X と同等な連続過程 Y (X の**連続変形**という) が存在し，しかもこの Y は強同等性を除いて一意である．

また Y はつぎの Hölder 連続性をもつ．

$$\lim_{h\to0}\sup_{\substack{0\leqq s<t\leqq a\\0<t-s<h}}\frac{|Y_t-Y_s|}{|t-s|^\varepsilon}=0 \text{ a.s.}$$

ここで ε は $(0,\gamma/\alpha)$ の中の任意の定数である．

証明　$X_t=X_t(\omega)$ を $X(t)=X(t,\omega)$ であらわすことにする．$X(t,\omega)$ のかわりに $X(at,\omega)$ を考えることにより，$a=1$ のばあいに帰着されるから，$a=1$ と仮定する．正の数 $\delta<\gamma/\alpha$ を任意にとり，これを固定する．Bienaymé の不等式

$$P\{|X|\geqq c\}\leqq c^{-\alpha}E(|X|^\alpha)$$

を用いて，仮定から

$$P\left\{\left|X\left(\frac{k}{2^n}\right)-X\left(\frac{k-1}{2^n}\right)\right|\geqq\left(\frac{1}{2^n}\right)^\delta\right\}\leqq 2^{n\delta\alpha}\cdot\beta2^{-n(1+\gamma)}=\beta2^{-n}2^{-n(\gamma-\alpha\delta)},$$

$$P\left(\bigcup_{k=1}^{2^n}\left\{\left|X\left(\frac{k}{2^n}\right)-X\left(\frac{k-1}{2^n}\right)\right|\geqq\left(\frac{1}{2^n}\right)^\delta\right\}\right)\leqq\beta2^{-n(\gamma-\alpha\delta)}.$$

δ のとり方により $\gamma-\alpha\delta>0$ であるから，右辺を項とする級数は収束し，Borel-Cantelli の補題により，つぎの事象の確率は 1 である．

$$\Omega_1=\left\{(\exists n_0=n_0(\omega))(\forall n\geqq n_0(\omega))\ \left|X\left(\frac{k}{2^n}\right)-X\left(\frac{k-1}{2^n}\right)\right|<2^{-n\delta}, k=1,2,\cdots,2^n\right\}.$$

さて $D=\{k/2^n\mid n=1,2,\cdots,k=0,1,2,\cdots,2^n\}$ は区間 $[0,1]$ の中で稠密である．$\omega\in\Omega_1$ に対し

(a)
$$|X(r,\omega)-X(r',\omega)|\leqq2(1-2^{-\delta})^{-1}(r'-r)^\delta,$$
$$0<r'-r\leqq2^{-n_0},\quad r,r'\in D$$

を証明しよう．これがいえると，$X(r,\omega)$ $(\omega\in\Omega_1)$ は $r\in D$ に関して一様連続となり，

$$Y(t,\omega)=\lim_{\substack{r\to t\\r\in D}}X(r,\omega),\qquad\omega\in\Omega_1$$

($\omega\in\Omega-\Omega_1$ では $Y(t,\omega)=0$ とおく)

が確定して，$\{Y_t\equiv Y(t,\omega)\}$ は連続過程となる．任意の $t\in[0,1]$ に対し，これに収束する列 $\{r_n\}\subset D$ をとると，

$$E(|Y_t-X_t|^\alpha)=E\left(\lim_{n\to\infty}|X(r_n)-X(t)|^\alpha\right)$$
$$\leqq\liminf_{n\to\infty}E(|X(r_n)-X(t)|^\alpha)\leqq\liminf_{n\to\infty}\beta|r_n-t|^{1+\gamma}=0$$

が得られ，$P\{Y_t=X_t\}=1$ となるから，$\{Y_t\}$ は $\{X_t\}$ と同等な連続過程となる．もし $\{Y_t'\}$ がこの性質をもつと，D の可算性により $P\{(\forall r\in D)\ Y_r=Y_r'\}=1$，$\{Y_t\},\{Y_t'\}$ の連続性により，$P\{(\forall t\in[0,1])\ Y_t=Y_t'\}=1$ が得られ，$\{Y_t\},\{Y_t'\}$ は強同等となる．

残る所は上の不等式 (a) の証明である．$0<r'-r\leqq2^{-n_0}$ により，

$$2^{-(n+1)}<r'-r\leqq2^{-n}$$

となる $n\geqq n_0$ が存在する．r,r' は同じ区間 $[l2^{-n},(l+1)2^{-n}]$ $(l=0,1,\cdots,2^n-1)$ の中に入るか，または隣接区間 $((l-1)2^{-n},l2^{-n}]$, $(l2^{-n},(l+1)2^{-n}]$ に入るかのいずれかである．

(i)　$l2^{-n}\leqq r<r'\leqq(l+1)2^{-n}$ のとき．$r-l2^{-n}$ を 2 進法展開して

$$r-l2^{-n}=\varepsilon_12^{-n-1}+\varepsilon_22^{-n-2}+\cdots+\varepsilon_p2^{-n-p}\qquad(\varepsilon_i=0,1)$$

とし，

$$r_0=l2^{-n},\qquad r_k=l2^{-n}+\varepsilon_12^{-n-1}+\cdots+\varepsilon_k2^{-n-k}\qquad(k=1,2,\cdots,p)$$

とおく．r_k は b_k2^{-n-k} $(b_k\in N)$ の形をしているから，$\omega\in\Omega_1$ に対しては

$$|X(r_k)-X(r_{k-1})|=|X(b_k2^{-n-k})-X(b_{k-1}2^{-n-k+1})|$$
$$=|X(b_k2^{-n-k})-X(2b_{k-1}2^{-n-k})|,$$
$$b_k-2b_{k-1}=2^{n+k}r_k-2\cdot2^{n+k-1}r_{k-1}=2^{n+k}(r_k-r_{k-1})=\varepsilon_k$$
$$=0 \text{ または } 1$$

であるから，Ω_1 の定義により

$$|X(r_k)-X(r_{k-1})|\leqq2^{-(n+k)\delta},$$

$$|X(r)-X(l2^{-n})|=|X(r_p)-X(r_0)|\leqq\sum_{k=1}^p|X(r_k)-X(r_{k-1})|$$

$$\leqq \sum_{k=1}^{p} 2^{-(n+k)\delta} < 2^{-(n+1)\delta}(1-2^{-\delta})^{-1}.$$

同様に $|X(r')-X(l2^{-n})|$ に対しても同じ評価を得るから

$$|X(r')-X(r)| < 2\cdot 2^{-(n+1)\delta}(1-2^{-\delta})^{-1}.$$

$r'-r>2^{-(n+1)}$ であるから，右辺は $2(1-2^{-\delta})^{-1}(r'-r)^{\delta}$ より小さい．

(ii) $(l-1)2^{-n}\leqq r\leqq l2^{-n}\leqq r'\leqq (l+1)2^{-n}$ のとき．$r'-r\leqq 2^{-n}$ により，上の条件の不等号'≦'の中等号がひとつでもあれば，(i) のばあいに帰着されるから，すべて'<'であるとしてよい．$r'-l2^{-n}$, $l2^{-n}-r$ はいずれも 2^{-n} より小さいから，これを2進法展開して，(i) と同様の論法を適用すれば，$|X(r')-X(r)|$ が $2(1-2^{-\delta})^{-1}(r'-r)^{\delta}$ で抑えられることがわかる．

これで定理の前半は証明された．

定理の後半をいうにも $a=1$ と仮定してよい．定数 $\varepsilon\in(0,\gamma/\alpha)$ に対し，$\varepsilon<\delta<\gamma/\alpha$ となる δ をとると，$\gamma-\alpha\delta>0$ であるから，上の証明にあるように

$$|X(r,\omega)-X(r',\omega)| \leqq c|r'-r|^{\delta} \qquad (|r'-r|<2^{-n_0(\omega)}) \text{ a.s.}$$

ここで r, r' は D の中を動き，c は $2(1-2^{-\delta})^{-1}$ である．$\{Y_t\}$ は $\{X_t\}$ と同等であり，D は可算であるから，

$$P\{Y_r=X_r,\ r\in D\}=1.$$

したがって上の式は Y に対してもなりたち，Y の連続性により，t, t' が $[0,1]$ の上を動くとしても

$$|Y(t,\omega)-Y(t',\omega)| \leqq c|t-t'|^{\delta}, \qquad |t'-t|<2^{-n_0(\omega)} \text{ a.s.}$$

がなりたつ．$\delta>\varepsilon$ であるから，これから定理の後半の結論がすぐでる．∎

さてもとの Wiener 過程の存在証明にもどろう．$X=\{X_t, t\in[0,\infty)\}$ を広義の Wiener 過程(この存在はすでに証明した)とする．$X_t-X_s\ (t>s)$ は $\mathcal{L}[X]$ に属するから，その確率法則は Gauss 分布 $N_{0,V}$ である．ただし V は

$$V=\|X_t-X_s\|^2=(X_t,X_t)-2(X_t,X_s)+(X_s,X_s)$$
$$=t\wedge t-2t\wedge s+s\wedge s=t-2s+s=t-s.$$

したがって $(X_t-X_s)/\sqrt{t-s}$ は $N_{0,1}$ に従い

$$E\{(X_t-X_s)^4\}=E\{(t-s)^2[(X_t-X_s)/\sqrt{t-s}]^4\}$$
$$=(t-s)^2\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}x^4dx$$

$$=3(t-s)^2 \qquad (\text{例題 } 2.6\text{ (ix)}).$$

したがって上述の定理の仮定が任意の a に対して成立している ($\alpha=4$, $\beta=3$, $\gamma=1$). これから任意の a に対し $X^{(a)}=\{X_t, t\in[0,a]\}$ は連続変形 $Y^{(a)}=\{Y_t^{(a)}, t\in[0,a]\}$ をもつ．$a<b$ のときには $Y^{(b)}$ を $[0,a]$ に制限したものは $X^{(a)}$ の連続変形となっているから，$Y^{(a)}$ と強同等である．したがって

$$P\{Y_t^{(n)}=Y_t^{(n+1)}\ (0\leqq t\leqq n),\ n=1,2,\cdots\}=1$$

となる．このような ω に対しては

$$B_t(\omega)=Y_t^{(n)} \quad (0\leqq t\leqq n), \qquad n=1,2,\cdots.$$

他の ω に対しては $B_t(\omega)=0$ とおくと，$B=\{B_t, t\in[0,\infty)\}$ は $X=\{X_t, t\in[0,\infty)\}$ の連続変形であることは明らかであろう．B は X と同等であるから，当然 (W.1), (W.2) をみたし，B は Wiener 過程である．

Wiener 過程は広義 Wiener 過程でもあるから，同等性を除いて一意であることは明らかであるが，連続過程でもあるから，さらに強同等を除いて一意である．

Wiener 過程の存在と一意性が確定したから，これから，Wiener 過程 $B=\{B(t), 0\leqq t<\infty\}$ の性質を調べよう．

(W.1) **加法性** $[s_1,t_1)$, $[s_2,t_2)$, … がたがいに素であるときには，$B(t_1)-B(s_1)$, $B(t_2)-B(s_2)$, … が独立である．

証明 $B=\{B(t)\}$ は平均0の Gauss 過程で，

$$B(t_i)-B(s_i)\subset\mathcal{L}[B]$$

であるから，$s\leqq t\leqq u\leqq v$ のとき，$B(t)-B(s)$ と $B(v)-B(u)$ とが $L^2(\Omega,P)$ の中で直交することをいえばよい．

$$(B(t)-B(s), B(v)-B(u))=t\wedge v-t\wedge u-s\wedge v+s\wedge u$$
$$=t-t-s+s=0. \qquad ∎$$

(W.2) **Wiener 過程を不変にする変換** Wiener 過程 B から得られる下の過程はいずれも Wiener 過程である:

$$B^1(t)=B(t+a)-B(a) \qquad (a \text{ は正の定数}),$$
$$B^2(t)=B(bt)/\sqrt{b} \qquad (b \text{ は正の定数}),$$
$$B^3(t)=-B(t),$$
$$B^4(t)=tB(1/t) \qquad \text{ただし} \quad B^4(0)=0.$$

証明 すべての t に対して $B^i(t)\subset\mathcal{L}[B]$ であるから，B^i は Gauss 過程であ

る．しかも

$$E(B^i(t))=0,\qquad V(B^i(t),B^i(s))=t\wedge s$$

となることは簡単な計算で示されるから，B^i は広義の Wiener 過程である．したがって B^i の連続性をいえば，B^i が Wiener 過程であることがわかる．B^i, $i=1,2,3$ の連続性は明らかである．また B^4 も $t>0$ で連続なことは明らかであるが，問題は $t=0$ での連続性：

$$P\left\{\lim_{t\to 0}B^4(t)=0\right\}=1$$

である．上に証明したことから，任意の n と任意の $t_1,t_2,\cdots,t_n\in(0,\infty)$ に対し，

$$P\{(B^4(t_1),B^4(t_2),\cdots,B^4(t_n))\in E_n\}=P\{(B(t_1),B(t_2),\cdots,B(t_n))\in E_n\},$$
$$E_n\in\mathcal{B}^n$$

がなりたつ．これから Dynkin 族定理を応用して

(a) $$P\{B_\cdot^4\in E\}=P\{B_\cdot\in E\},\qquad E\in\mathcal{B}_K(\boldsymbol{R}^{[0,\infty)})$$

が得られる．$B_\cdot^4$ は開区間 $(0,\infty)$ では連続であるから，

$$\left\{\lim_{t\to 0}B^4(t)=0\right\}=\left\{B_\cdot^4\in\bigcap_m\bigcup_n\bigcap_{0<r<1/n}\pi_r^{-1}\left(-\frac{1}{m},\frac{1}{m}\right)\right\}.$$

ここで r は有理数のみを動き，π_r は射影

$$\boldsymbol{R}^{[0,\infty)}\longrightarrow\boldsymbol{R}^1,\qquad f\longmapsto f(r)$$

である．B についても同じことがいえるから，(a) により

$$P\left\{\lim_{t\to 0}B^4(t)=0\right\}=P\left\{\lim_{t\to 0}B(t)=0\right\}=1$$

が得られる．▌

上に証明した

$$\lim_{t\to 0}B^4(t)=0\ \text{a.s.}\quad すなわち\quad \lim_{t\to 0}tB\left(\frac{1}{t}\right)=0\ \text{a.s.}$$

において t を $1/t$ でおきかえると

$$\lim_{t\to\infty}\frac{B(t)}{t}=0\ \text{a.s.}$$

が得られる．この事実は t が整数だけ動くときには，

$$X_n=B(n)-B(n-1),\qquad n=1,2,\cdots$$

に大数の強法則を適用しても得られる．しかし t が連続的に動くときには $B(t+n)-B(n)$, $0\leqq t\leqq 1$ を評価する必要があり，もう少し議論を要する．

(W. 3) **Wiener 過程の Hölder 一様連続性** (Paley-Wiener-Zygmund)

$0<\varepsilon<1/2$ か $\varepsilon\geqq 1/2$ に応じて

$$\lim_{h\to 0}\sup_{\substack{|t-s|<h\\ 0\leqq t,s\leqq a}}\frac{|B(t)-B(s)|}{|t-s|^\varepsilon}=0\ \text{または}\ \infty\quad\text{a.s.}$$

証明 $B(at)/\sqrt{a}$ も Wiener 過程であるから，$a=1$ のばあいを考察すれば十分である．$(B_t-B_s)/\sqrt{t-s}$ は $N_{0,1}$ に従うから，

$$E\left\{\frac{(B_t-B_s)^{2n}}{(t-s)^n}\right\}=c_n\equiv 1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)\qquad(\text{例題 2. 6 (ix)}).$$

ゆえに

$$E((B_t-B_s)^{2n})=c_n|t-s|^n=c_n|t-s|^{1+(n-1)}.$$

定理 5. 24 の後半により，任意の $0<\varepsilon<(n-1)/2n$ に対して上の極限は 0 になる．n を大きくとることにより，$(n-1)/2n$ はいくらでも 1/2 に近くできるから，$0<\varepsilon<1/2$ に対しても上の極限は 0 になる．

$\varepsilon\geqq 1/2$ のときを考えよう．$\varepsilon=1/2$ のばあいについて上の極限が ∞ になることをいえば十分である．Wiener 過程の加法性により

$$P\left(\bigcap_{k=1}^n\left\{\left|B\left(\frac{k}{n}\right)-B\left(\frac{k-1}{n}\right)\right|\leqq a\sqrt{\frac{1}{n}}\right\}\right)$$
$$=\prod_{k=1}^n P\left\{\left|B\left(\frac{k}{n}\right)-B\left(\frac{k-1}{n}\right)\right|\leqq a\sqrt{\frac{1}{n}}\right\}.$$

右辺の因子はすべて $\Phi(a)\equiv N_{0,1}[-a,a]$ に等しいから，左辺の事象 (ω 集合) を Ω_n であらわすと，

$$\sum_n P(\Omega_n)=\sum_n\Phi(a)^n<\infty.$$

Borel-Cantelli の補題により，ほとんど確実に，或る番号 $n_0(\omega)$ 以後事象 Ω_n が起らない．すなわち $n>n_0$ に対して

$$\left|B\left(\frac{k}{n}\right)-B\left(\frac{k-1}{n}\right)\right|>a\sqrt{\frac{1}{n}}=a\left(\frac{k}{n}-\frac{k-1}{n}\right)^{1/2}$$

がなりたつような $k\ (\leqq n)$ がある．これから問題の極限 l はほとんど確実に a 以上である．$a\to\infty$ として $l=\infty$ a.s. を得る．▌

上の事実から Wiener 過程の見本過程の微分不可能な点が稠密であることは容

易に証明されるが，さらに進んで，つぎの事実が得られる．

(W. 4)　Wiener 過程の見本過程はほとんど確実に到るところ微分不可能である．(Paley-Wiener-Zygmund)

証明　Dvoretzki-Erdös-Kakutani による簡単な証明を紹介する．

$$A = \{B_t(\omega) \text{ が } [0,1] \text{ の中のある点で微分可能}\}$$

とおいて，$P(A)=0$ を証明すればよいことは容易にわかる．

$$A_l = \left\{(\exists t \in [0,1]) \ \left|\frac{dB_t}{dt}(\omega)\right| < l\right\}$$

とおくと，$A=\bigcup_l A_l$ であるから，$P(A_l)=0$ をいえばよい．

$\omega \in A_l$ ならば，ある $t_0=t_0(\omega) \in [0,1]$ があって $dB_t(\omega)/dt|_{t=t_0}$ の絶対値が l より小さいから，$\delta=\delta(t_0, \omega)$ が存在して，

$$t_0-\delta < s < t_0 < t < t_0+\delta \Longrightarrow |B(t)-B(s)| < l(t-s).$$

さて $N=N(\omega)$ を十分大きくとれば，$n>N$ なる限り

$$t_0-\delta \leqq \frac{k}{n} \leqq t_0 < \frac{k+1}{n} < \frac{k+2}{n} < \frac{k+3}{n} < t_0+\delta$$

となる $k=k(\omega) \leqq n$ がある．この k に対しては

$$\left|B\left(\frac{k+i}{n}, \omega\right)-B\left(\frac{k}{n}, \omega\right)\right| < \frac{li}{n}, \qquad i=1,2,3,$$

したがって

$$\left|B\left(\frac{k+1}{n}, \omega\right)-B\left(\frac{k}{n}, \omega\right)\right| \leqq \frac{l}{n}, \qquad \left|B\left(\frac{k+2}{n}, \omega\right)-B\left(\frac{k+1}{n}, \omega\right)\right| \leqq \frac{3l}{n},$$
$$\left|B\left(\frac{k+3}{n}, \omega\right)-B\left(\frac{k+2}{n}, \omega\right)\right| \leqq \frac{5l}{n}$$

がなりたつ．

さて一般の n, k に対して，上の条件をみたす ω の集合を A_{nk} であらわすと，上の考察から

$$A_l \subset \bigcup_N \bigcap_{n>N} \bigcup_{k=1}^{n} A_{nk}.$$

右辺の P 測度が 0 であることをいえば，P の完備性により，$P(A_l)=0$ となる．

$$P(A_{nk}) = \prod_{i=1}^{3} P\left\{\left|B\left(\frac{k+i}{n}\right)-B\left(\frac{k+i-1}{n}\right)\right| \leqq \frac{(2i-1)l}{n}\right\},$$

$$P\{|B(t)-B(s)| \leqq c(t-s)\} \leqq \int_{-c(t-s)}^{c(t-s)} \frac{dx}{\sqrt{2\pi(t-s)}} = c\sqrt{\frac{2}{\pi}(t-s)}$$

により

$$P(A_{nk}) \leqq c'\left(\frac{1}{n}\right)^{3/2} \qquad (c' \text{ は } n, k \text{ に無関係な定数}),$$

$$P\left(\bigcup_{k=1}^{n} A_{nk}\right) \leqq \sum_{k=1}^{n} P(A_{nk}) \leqq c'\left(\frac{1}{n}\right)^{1/2} \longrightarrow 0 \qquad (n \to \infty),$$

$$P\left(\bigcap_{n>N} \bigcup_{k=1}^{n} A_{nk}\right) = 0, \qquad \text{ゆえに} \quad P\left(\bigcup_N \bigcap_{n>N} \bigcup_{k=1}^{n} A_{nk}\right) = 0. \qquad \blacksquare$$

Wiener 過程 $B=\{B_t, t \in [0,\infty)\}$ で生成される増大情報系

$$\boldsymbol{F}[B] = \{\mathcal{F}_t[B], t \in [0,\infty)\}$$

を考えよう．$\mathcal{F}_t[B]$ の定義により

$$\mathcal{F}_t[B] = \bigcap_n \sigma\left[B_s, s \leqq t+\frac{1}{n}\right] \vee \boldsymbol{2}.$$

(W. 5)　$\mathcal{F}_t[B]$ と $\sigma[B_{s+t}-B_t, s \in [0,\infty)]$ とは独立である．

証明　加法性により

$$0 \leqq t_1 < t_2 < \cdots < t_m \leqq t, \qquad 0 \leqq s_1 < s_2 < \cdots < s_n$$

のとき

$$\sigma[B(t_1), B(t_2)-B(t_1), \cdots, B(t_m)-B(t_{m-1})]$$

と

$$\sigma[B(t+s_1)-B(t), B(t+s_2)-B(t+s_1), \cdots, B(t+s_n)-B(t+s_{n-1})]$$

とは独立である．これらはそれぞれ

$$\sigma[B(t_1), B(t_2), \cdots, B(t_m)],$$
$$\sigma[B(t+s_1)-B(t), B(t+s_2)-B(t), \cdots, B(t+s_n)-B(t)]$$

に等しい．このおのおのに $\bigvee_{t_1, t_2, \cdots, t_m}$ $(m=1,2,\cdots, t_i \in [0,t])$，$\bigvee_{s_1, s_2, \cdots, s_n}$ $(n=1,2,\cdots, s_i \in [0,\infty))$ を施して

$$\sigma[B(u), u \leqq t] \quad \text{と} \quad \sigma[B(s+t)-B(t), s \in [0,\infty)]$$

とが独立であることがわかる．したがって

$$\sigma\left[B(u), u \leqq t+\frac{1}{n}\right] \quad \text{と} \quad \sigma\left[B\left(s+t+\frac{1}{n}\right)-B\left(t+\frac{1}{n}\right), s \in [0,\infty)\right]$$

とも独立である．これから

$$\mathscr{F}_t[B] \quad と \quad \sigma\left[B\left(s+t+\frac{1}{n}\right)-B\left(t+\frac{1}{n}\right), s\in[0,\infty)\right]$$

とはすべての $n=1,2,\cdots$ に対して独立であるから，

$$\mathscr{F}_t[B] \quad と \quad \bigvee_n \sigma\left[B\left(s+t+\frac{1}{n}\right)-B\left(t+\frac{1}{n}\right), s\in[0,\infty)\right]$$

とも独立である．$B(t)$ は t について連続であるから，右側の σ 加法族は

$$\sigma[B(s+t)-B(t), s\in[0,\infty)]$$

と一致する．これで (W.5) は証明された．▌

(W.6) τ を $\boldsymbol{F}[B]$ 停止時とすると，

$$P\{\tau\leqq t, B_t-B_\tau\in E\}=\int_{[0,t]}P^\tau(ds)P\{B_t-B_s\in E\}, \qquad E\in B^1.$$

ここで P^τ は τ の確率法則である．さらに一般に

$$P\{B_t-B_{\tau\wedge t}\in E\mid \mathscr{F}_{\tau\wedge t}[B]\}=P\{B_t-B_s\in E\}|_{s=\tau\wedge t}.$$

証明 $\{\tau\leqq t\}\in\mathscr{F}_{\tau\wedge t}[B]$ であるから，第1式は第2式の特別のばあいである．

$$\tau_n=\frac{[n\tau]+1}{n}\wedge t, \qquad n=1,2,\cdots$$

とおく．任意の有界連続関数 $f:\boldsymbol{R}^1\to\boldsymbol{R}^1$ と任意の $A\in\mathscr{F}_\tau[B]$ に対し

$$\begin{aligned}E(f(B_t-B_{\tau_n}),A)&=\sum_{k=1}^\infty E\left(f(B_t-B_{k/n\wedge t}), A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right)\\&=\sum_{k=1}^{[nt]}E\left(f(B_t-B_{k/n}), A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right)\\&\quad+\sum_{k>[nt]}f(0)P\left(A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right).\end{aligned}$$

A が $\mathscr{F}_\tau[B]$ に属するから，$k\leqq n$ のときには

$$A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}=A\cap\left\{\tau<\frac{k}{n}\right\}-A\cap\left\{\tau<\frac{k-1}{n}\right\}\in\mathscr{F}_{k/n}[B].$$

したがって (W.5) により，上の式は

$$=\sum_{k=1}^\infty E(f(B_t-B_{k/n\wedge t}))P\left(A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right),$$

また上と同様の計算により

$$E(E(f(B_t-B_s))|_{s=\tau_n},A)=\sum_{k=1}^\infty E(f(B_t-B_{k/n\wedge t}))P\left(A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right).$$

これから

$$E(f(B_t-B_{\tau_n}),A)=E(E(f(B_t-B_s))|_{s=\tau_n},A).$$

B_t は s について連続であるから，有界収束定理により，$E(f(B_t-B_s))$ は s について連続．したがって $\tau_n\to\tau\wedge t$ に注意して

$$E(f(B_t-B_{\tau\wedge t}),A)=E(E(f(B_t-B_s))|_{s=\tau\wedge t},A)$$

が得られる．これから常套の極限論法で，$f=1_E$ $(E\in\mathscr{B}^1)$ のときにも上の式がなりたつことがわかり，任意の $A\in\mathscr{F}_\tau[B]$ に対し

$$P\{\{B_t-B_\tau\in E\}\cap A\}=E(P\{B_t-B_s\in E\}|_{s=\tau\wedge t},A), \qquad E\in\mathscr{B}^1.$$

すなわち (W.6) の第2式がなりたつ．▌

(W.2) と (W.5) とからわかるように，

$$B^t=\{B(s+t)-B(t), s\in[0,\infty)\}$$

は $\mathscr{F}_t[B]$ と独立な Wiener 過程である．この事実はつぎのように一般化される．

(W.7) τ が(ほとんど確実に)有限な $\boldsymbol{F}[B]$ 停止時であるときには

$$B^\tau=\{B(s+\tau)-B(\tau), s\in[0,\infty)\}$$

は $\mathscr{F}_\tau[B]$ と独立な Wiener 過程である．

証明 $\tau_n=([n\tau]+1)/n$ とおくと，

$$\tau_n=\frac{k}{n}\iff\frac{k-1}{n}\leqq\tau<\frac{k}{n} \quad かつ \quad \tau_n\to\tau\ (n\to\infty)$$

となる．$B_n(s)=B(s+\tau_n)-B(\tau_n)$ とおいて，$B_n=\{B_n(s), s\in[0,\infty)\}$ が $\mathscr{F}_\tau[B]$ と独立な Wiener 過程であることをいう．$E_m\in\mathscr{B}^m$，$A\in\mathscr{F}_\tau[B]$ に対し

$$\begin{aligned}&P(\{(B_n(s_1),\cdots,B_n(s_m))\in E_m\}\cap A)\\&=\sum_{k=1}^\infty P\left(\{(B_n(s_1),\cdots,B_n(s_m))\in E_m\}\cap A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right)\\&=\sum_{k=1}^\infty P\left(\left\{\left(B\left(s_1+\frac{k}{n}\right)-B\left(\frac{k}{n}\right),\cdots,B\left(s_m+\frac{k}{n}\right)-B\left(\frac{k}{n}\right)\right)\in E_m\right\}\cap A\right.\\&\qquad\left.\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right).\end{aligned}$$

(W.5) の証明と同様に

$$=\sum_{k=1}^\infty P\left\{\left(B\left(s_1+\frac{k}{n}\right)-B\left(\frac{k}{n}\right),\cdots,B\left(s_m+\frac{k}{n}\right)-B\left(\frac{k}{n}\right)\right)\in E_m\right\}$$

$$\cdot P\left(A\cap\left\{\frac{k-1}{n}\leqq\tau<\frac{k}{n}\right\}\right)$$

$$=P\{(B(s_1),\cdots,B(s_m))\in E_m\}P(A)\qquad ((\mathrm{W}.2)\text{ による}).$$

とくに $A=\Omega$ とおくと,

$$P\{(B_n(s_1),\cdots,B_n(s_m))\in E_m\}=P\{(B(s_1),\cdots,B(s_m))\in E_m\}.$$

これを上の式にいれて

$$P(\{(B_n(s_1),\cdots,B_n(s_m))\in E_m\}\cap A)=P\{(B_n(s_1),\cdots,B_n(s_m))\in E_m\}P(A).$$

B_n は明らかに連続過程であるから，上の第1式から，B_n が Wiener 過程であることがわかり，第2式から B_n と $\mathcal{F}_\tau[B]$ との独立性が得られる．すべての t に対し $B_n(t)\to B(t)$ で，しかも $B(t)$ が連続過程であるから，B も $\mathcal{F}_\tau[B]$ と独立な Wiener 過程である．▮

(W. 8)　Wiener 過程 B は $\boldsymbol{F}[B]$ に関してマルチンゲールである．

証明　(W. 7) により，$t>s$ に対し

$$E(B_t-B_s\,|\,\mathcal{F}_s[B])=E(B_t-B_s)=0.$$

B_s は $\mathcal{F}_s[B]$ に関して可測であるから，これは

$$E(B_t\,|\,\mathcal{F}_s[B])=B_s\qquad (t>s)$$

を意味する．▮

下に示す Wiener 過程の特長づけは極めて有用である．

(W. 9)　$X=\{X_t, t\in[0,\infty)\}$ が Wiener 過程であるための必要十分条件はつぎの3条件がなりたつことである．

(i)　X は加法性をもつ．

(ii)　$X_0\equiv 0$ で，$X_t-X_s\ (t>s)$ は $N_{0,t-s}$ に従う．

(iii)　X は連続過程である．

証明　Wiener 過程がこの3条件をみたすことは明らかである．逆に X がこの3条件をみたすとせよ．まず X が平均0の Gauss 過程であることをいう．このためには $0=t_0<t_1<\cdots<t_n$ に対し

$$Y=\sum_{i=1}^{n}a_iX_{t_i}\qquad (a_i\text{ は実定数})$$

が平均0の Gauss 分布に従うことをいえばよい．Y は

$$Y=\sum_{i=1}^{n}b_i(X_{t_i}-X_{t_{i-1}})\qquad (X_0\equiv 0\text{ に注意せよ})$$

の形にかけ，(i), (ii) により $X_{t_i}-X_{t_{i-1}}$ は独立で $N_{0,t_i-t_{i-1}}$ に従うから，Y は平均0の Gauss 分布に従う．

また $s<t$ のとき，加法性により

$$\begin{aligned}E(X_tX_s)&=E((X_t-X_s)X_s)+E(X_s^2)\\&=E(X_t-X_s)E(X_s)+E(X_s^2)\\&=0+s=s\end{aligned}$$

であるから，一般に

$$E(X_tX_s)=t\wedge s.$$

したがって X は広義の Wiener 過程である．ゆえに (iii) により Wiener 過程である．▮

$\boldsymbol{F}=\{\mathcal{F}_t, t\in[0,\infty)\}$ を増大情報系とする．$X=\{X_t, t\in[0,\infty)\}$ を $\boldsymbol{F}$ に適合する連続過程で，

(i)　$X_0\equiv 0$,

(ii)　$X_t-X_s\ (t>s)$ は $\mathcal{F}_s$ と独立で，$N_{0,t-s}$ に従う

と仮定する．これから $0=t_0<t_1<\cdots<t_n$ のとき，

$$\begin{aligned}&P\{X_{t_i}-X_{t_{i-1}}\in E_i, i=1,2,\cdots,n\}\\&\quad=P\{X_{t_i}-X_{t_{i-1}}\in E_i, i=1,2,\cdots,n-1\}P\{X_{t_n}-X_{t_{n-1}}\in E_n\}\\&\qquad(\because\ \{X_{t_i}-X_{t_{i-1}}\in E_i, i=1,2,\cdots,n-1\}\in\mathcal{F}_{t_{n-1}}),\end{aligned}$$

これをくりかえして，

$$=\prod_{i=1}^{n}P\{X_{t_i}-X_{t_{i-1}}\in E_i\}$$

がでる．これは X が加法性をもつことを意味する．(W. 9) の他の条件は当然なりたつから，X が Wiener 過程であることがわかる．このような Wiener 過程 X を **$\boldsymbol{F}$ に従属する Wiener 過程**という．また $X_t-X_s\ (t>s)$ は $\mathcal{F}_s$ と独立であるから，

$$E(X_t-X_s\,|\,\mathcal{F}_s)=E(X_t-X_s)=0.$$

X_s は可測 $\mathcal{F}_s$ であるから，これから

$$E(X_t\,|\,\mathcal{F}_s)=X_s$$

がでる．これは X が F マルチンゲールであることを意味する．この意味で X を F に関する **Wiener マルチンゲール**ということがある．

前述の Wiener 過程 B は $F[B]$ に関する Wiener マルチンゲールであることは明らかである．(W. 6), (W. 7) は $F[B]$ を F でおきかえることにより，F に関する Wiener マルチンゲールに対してもなりたつ．

$X=\{X_t, t\in[0,\infty)\}$ を連続過程とし，$X.(\omega)$ をその見本過程とする．$X.(\omega)$ は C 値位相確率変数 $(C=C[0,\infty))$ である．これは $\mathcal{B}_K(C)=\mathcal{B}(C)$ からわかる．したがって $X.(\omega)$ の確率法則 $P^{X.}$ は C の上の正則確率測度であって，その行動は $\mathcal{B}(C)$ すなわち $\mathcal{B}_K(C)$ の上の行動で完全にきまる．このことに注意すると，$P^{X.}$ は $\{X_t\}$ の有限個の時点の上の結合分布を知ることにより，完全にきまる．

Wiener 過程は連続で，時点 $t_1, t_2, \cdots, t_n$ における結合分布は平均値ベクトル $(0,0,\cdots,0)$，分散行列 $(t_i\wedge t_j)_{i,j=1}{}^n$ の Gauss 分布であるから，すべての Wiener 過程の見本過程の確率法則は同じで，これを **Wiener 測度**という．これは $C[0,\infty)$ の上の正則確率測度である．

連続過程 $X=\{X_t, t\in[0,\infty)\}$, $Y=\{Y_t, t\in[0,\infty)\}$ が独立であるとは $\bigvee_t \sigma[X_t]$ と $\bigvee_t \sigma[Y_t]$ とが独立なことである．$\mathcal{B}_K(C)=\mathcal{B}(C)$ に注意すると，この σ 加法族はそれぞれ $X.^{-1}(\mathcal{B}(C))$, $Y.^{-1}(\mathcal{B}(C))$ に等しいから，X, Y の独立性はその見本過程 $X., Y.$ の独立性と同等である．

連続過程見本過程に関する上述の事実は D 過程に関してもなりたつことは $\mathcal{B}_K(D)=\mathcal{B}(D)$ から明らかである．

例題 5.8 $B=\{B_t, t\in[0,\infty)\}$ を Wiener 過程とする．

(i) (**Wiener 過程に関する反射定理**) τ を $F[B]=\{\mathcal{F}_t[B]\}$ に関する停止時とするとき，

$$\tilde{B}(t)=\begin{cases} B(t), & t\leqq\tau, \\ B(\tau)-(B(t)-B(\tau))=2B(\tau)-B(t), & t>\tau \end{cases}$$

も Wiener 過程であることを証明せよ．

[ヒント] $$B^0(t)=B(t\wedge\tau), \qquad \tau\in[0,\infty)$$

とおくと，固定した t に対して，$t\wedge\tau$ も $F[B]$ 停止時であるから，$B^0(t)$ は可測 $\mathcal{F}_{t\wedge\tau}[B]$ である．$\mathcal{F}_{t\wedge\tau}[B]\subset\mathcal{F}_\tau[B]$ により，$B^0(t)$ は可測 $\mathcal{F}_\tau[B]$ である．τ は可測 $\mathcal{F}_\tau[B]$ であるから，結合変数 $(B.^0,\tau)$ ($B.^0$ は B^0 の見本過程) も可測 $\mathcal{F}_\tau[B]$ である．

$$B^1(t)=B(t+\tau)-B(\tau), \qquad t\in[0,\infty)$$

は $\mathcal{F}_\tau[B]$ と独立な Wiener 過程である．同様に

$$B^2(t)=-(B(t+\tau)-B(\tau)), \qquad t\in[0,\infty)$$

も $\mathcal{F}_\tau[B]$ と独立な Wiener 過程である．したがって見本過程 $B.^1, B.^2$ はともに $(B.^0,\tau)$ と独立な Wiener 過程であり，したがって結合変数 $(B.^0,\tau,B.^1)$, $(B.^0,\tau,B.^2)$ は同じ確率法則をもつ．さて

$$\Phi:\ C\times[0,\infty)\times C\longrightarrow C$$

$$(f,s,g)\longmapsto F(t)=\begin{cases} f(t), & t\leqq s, \\ f(s)+g(t-s), & t>s \end{cases}$$

と定義すると，Φ は Borel 可測である．明らかに

$$B.=\Phi(B.^0,\tau,B.^1), \qquad \tilde{B}.=\Phi(B.^0,\tau,B.^2)$$

であるから，$\tilde{B}.$ と $B.$ とは同じ確率法則をもつ．ゆえに $\{\tilde{B}_t\}$ は Wiener 過程である．

(ii) $M(t)=\max\{B(s)\,|\,0\leqq s\leqq t\}$, $\quad \tau=\tau(a)=\inf\{t\,|\,B(t)\geqq a\}$ $\quad (a\geqq 0)$ のとき

$$P\{M(t)\geqq a\}=P\{\tau(a)\leqq t\}=2P\{B(t)\geqq a\}=P\{|B(t)|\geqq a\}.$$

$M(t), \tau(a)$ の確率法則は密度をもち

$$P\{M(t)\in da\}=\sqrt{\frac{2}{\pi t}}e^{-a^2/2t}da \qquad (a\geqq 0),$$

$$P\{\tau\in dt\}=\frac{1}{\sqrt{2\pi t^3}}e^{-a^2/2t}a\,dt \qquad (t\geqq 0).$$

[ヒント]
$$\begin{aligned} B(t)\geqq a &\Longrightarrow M(t)\geqq a,\quad B(t)\geqq a \\ &\Longrightarrow M(t)\geqq a,\quad B(t)-B(\tau)\geqq 0 \\ &\Longrightarrow \tau\leqq t,\quad B(t)-B(\tau)\geqq 0. \end{aligned}$$

したがって (W. 6) を用いて

$$\begin{aligned} P(B(t)\geqq a)&=P\{\tau\leqq t,\ B(t)-B(\tau)\geqq 0\} \\ &=\int_{[0,t]}P^\tau(ds)P(B_t-B_s\geqq 0)=\frac{1}{2}P^\tau[0,t]=\frac{1}{2}P\{\tau\leqq t\}. \end{aligned}$$

これから

$$P\{M(t)\geqq a\} = P\{\tau(a)\leqq t\} = 2P\{B(t)\geqq a\} = P\{|B(t)|\geqq a\}$$

が得られる．(最後の等式は $B(t)$ の分布の対称性による．) ゆえに

$$P\{M(t)\geqq a\} = P\{\tau(a)\leqq t\} = \sqrt{\frac{2}{\pi t}}\int_a^\infty e^{-x^2/2t}dx = \sqrt{\frac{2}{\pi}}\int_{a/\sqrt{t}}^\infty e^{-x^2/2}dx.$$

これを a または t で微分して，$M(t), \tau(a)$ の分布密度が得られる．

(iii)　$[a, b]$ を有限区間とし，x をその中の点とする．

$$B^x(t) = x+B(t), \qquad t\in[0,\infty)$$

を x から出発する Wiener 過程という．$\tau=\tau(x, a, b, \omega)$ を B^x が $[a, b]$ の端点(a または b) に達する時点(簡単に (a, b) からでる時点)という．このとき $P(\tau<\infty)=1$ を証明せよ．さらに

$$p_a = P\{B^x(\tau)=a\}, \qquad p_b = P\{B^x(\tau)=b\}, \qquad m = E\{\tau\}$$

を求めよ．p_a (p_b) は B^x が始めて (a, b) からでるとき，a (b) からでる確率をあらわす．

［ヒント］　$P\{\tau=\infty\} \leqq P\{B^x(t)\in(a,b)\} = P\{B(t)\in(a-x, b-x)\}$

$$= P\{B^x(1)\in((a-x)/\sqrt{t}, (b-x)/\sqrt{t})\} \longrightarrow 0 \qquad (t\to\infty)$$

により，$P\{\tau<\infty\}=1$ が得られる．B が $\boldsymbol{F}[B]$ に関してマルチンゲールであるから，B^x も同様である．$B^x(t\wedge\tau)$ は $[a, b]$ に属するから，$B^x(t\wedge\tau)$, $t\in[0,\infty)$ は有界な $\boldsymbol{F}[B]$ マルチンゲールである．これから

$$E(B^x(\tau)) = E\left(\lim_{t\to\infty} B^x(t\wedge\tau)\right) = E(B^x(0)) = x \qquad (\text{定理 } 5.17),$$

すなわち $ap_a+bp_b=x$．これと $p_a+p_b=1$ とを用いて

$$p_a = \frac{b-x}{b-a}, \qquad p_b = \frac{x-a}{b-a}$$

を得る．また

$$X(t) = B(t)^2-t$$

も $\boldsymbol{F}[B]$ マルチンゲールである．実際 $t>s$ のとき (W.5) により ($\mathcal{F}_s[B]$ を $\mathcal{F}_s$ とかく)

$$\begin{aligned} E(B_t^2\,|\,\mathcal{F}_s) &= B_s^2+2B_sE(B_t-B_s\,|\,\mathcal{F}_s)+E((B_t-B_s)^2\,|\,\mathcal{F}_s) \\ &= B_s^2+2B_sE(B_t-B_s)+E((B_t-B_s)^2) \end{aligned}$$

$$= B_s^2+t-s$$

より $E(X_t\,|\,\mathcal{F}_s)=X_s$ がでる．したがって $X(t\wedge\tau\wedge n)$, $t\in[0,\infty)$ は有界な $\boldsymbol{F}[B]$ マルチンゲールである．

$$E\left(\lim_{t\to\infty} X(t\wedge\tau\wedge n)\right) = E(X(0)) = x^2,$$

すなわち

$$E(B(\tau\wedge n)^2)-E(\tau\wedge n) = x^2.$$

有界収束定理と単調収束定理により，$n\to\infty$ として

$$E(B(\tau)^2)-E(\tau) = x^2.$$

これから $E(\tau)<\infty$ がわかり

$$E(B(\tau)^2-\tau) = x^2.$$

これから

$$\begin{aligned} E(\tau) &= E(B(\tau)^2)-x^2 = a^2P(B(\tau)=a)+b^2P(B(\tau)=b)-x^2 \\ &= a^2\frac{b-x}{b-a}+b^2\frac{x-a}{b-a}-x^2 = -(b-x)(a-x). \end{aligned}$$

(iv)　B^1, B^2 が独立な Wiener 過程であるとき

$$B(t) = a_1B^1(t)+a_2B^2(t) \qquad (a_1, a_2 \text{ は実定数で } a_1{}^2+a_2{}^2=1 \text{ とする})$$

も Wiener 過程であることを示せ．

［ヒント］　$\mathcal{B}_t = \sigma[B_s^1, B_s^2, s\leqq t]$,　$\mathcal{F}_t = \bigwedge_n \mathcal{B}_{t+1/n} \vee \mathbf{2}$

とおくと，情報系 $\boldsymbol{F}=\{\mathcal{F}_t, t\in[0,\infty)\}$ を得る．B^1, B^2 がいずれも $\boldsymbol{F}$ に関する Wiener マルチンゲールであることを示せば，B も同様であることがわかる．

§5.9　多項配置，Poisson 配置

S を一般の集合とし，$\mathcal{S}$ を S の上の σ 加法族とする．$\mathcal{S}$ について

($\mathcal{S}$.1)　すべての $x\in S$ に対し $\{x\}\in\mathcal{S}$,

($\mathcal{S}$.2)　$\mathcal{S}$ は可算個の集合で生成される

の2条件をつけておく．このとき S に $\mathcal{S}$ をそえたもの $(S, \mathcal{S})$ を **Borel 空間**とよぶ．

S の上の測度 μ の中でつぎの条件をみたすもののみを考え，これを簡単に $(S, \mathcal{S})$ の上の測度という．

$(\mu.1)$ $\mathcal{D}(\mu)\supset\mathcal{S}$,

$(\mu.2)$ (σ 有限性)

$$S=\bigcup_n E_n \qquad (E_n\in\mathcal{S},\ \mu(E_n)<\infty),$$

$(\mu.3)$ (**完全性**) 任意の $E\in\mathcal{S}$ $(0<\mu(E)<\infty)$ に対し, $\nu_E(A)\equiv\mu(A)/\mu(E)$, $A\in\mathcal{D}(\mu)\cap E$, は E の上の完全確率測度である.

$(\mathcal{S}.1)$ により, $\mathcal{S}$ は分離族であり, したがって $\mathcal{S}$ を生成する可算族も分離族である. これから, $(\mu.3)$ の ν_E は可分完全となり, 定理 3.1* を利用して, ν_E がその制限 $\nu_E|_{\mathcal{S}\cap E}$ の Lebesgue 拡大であることがわかる. したがって, $(\mu.2)$ により, μ も $\mu|_{\mathcal{S}}$ の Lebesgue 拡大となり μ はその $\mathcal{S}$ の上の行動で完全に決定される.

$(S,\mathcal{S})$ の上の測度 μ が**連続**であるとは, すべての $x\in S$ に対し, $\mu(\{x\})=0$ となることである.

$(\boldsymbol{R}^n,\mathcal{B}^n)$ $(n=1,2,\cdots,\infty)$ は Borel 空間である. 一般に S が可算開基をもつ位相空間であるとき, $(S,\mathcal{B}(S))$ は Borel 空間である. S が完備可分距離空間のとき, S の上の σ 有限な正則測度 μ (μ が $\mu|_{\mathcal{B}(S)}$ の Lebesgue 拡大であるとき, 確率測度のばあいと同様に μ を**正則**という) は上の意味で $(S,\mathcal{B}(S))$ の上の測度である. $\boldsymbol{R}^n$ $(n<\infty)$ の上の Lebesgue 測度は上の意味で $(\boldsymbol{R}^n,\mathcal{B}^n)$ の上の連続な測度である.

μ を Borel 空間 $(S,\mathcal{S})$ の上の連続な測度とするとき, 任意の $E\in\mathcal{S}$ $(0<\mu(E)<\infty)$ に対し, つぎの事実がなりたつことを証明しておこう.

(i) 任意の $\varepsilon>0$ に対し, $E'\subset E$ があって $0<\mu(E')<\varepsilon$.

(ii) 任意の $\alpha\in[0,1]$ に対し, $E'\subset E$ があって $\mu(E')=\alpha\mu(E)$.

(iii) 任意の $\varepsilon>0$ に対し, E の分割

$$E=\sum_{k=1}^n E_k, \qquad \mu(E_k)<\varepsilon \quad (k=1,2,\cdots,n)$$

がある.

[(i) の証明] $\mathcal{S}$ を生成する可算族を $\{A_n\}$ とする. $\mu(E)>0$ により $E\cap A_1$, $E\cap A_1{}^c$ の一方は正の μ 測度をもつ. これを B_1 とすると, $B_1\cap A_2$, $B_1\cap A_2{}^c$ の一方は正の μ 測度をもつ. これを B_2 とする. これをくりかえして $B_3, B_4, \cdots$ を定義すると

$$B_1\supset B_2\supset\cdots, \qquad \mu(B_n)>0.$$

しかも

$$\bigcap_n B_n=E\cap A_1{}'\cap A_2{}'\cap\cdots, \qquad A_n{}'=A_n \text{ または } A_n{}^c$$

である. $\{A_n\}$ が $\mathcal{S}$ を生成することから, $\bigcap_n B_n$ が空集合または1点集合であることがわかる. μ の連続性により,

$$\lim_{n\to\infty}\mu(B_n)=\mu\Big(\bigcap_n B_n\Big)=0.$$

したがって, n を十分大きくとると $0<\mu(B_n)<\varepsilon$. この B_n を E' とおけばよい. ▌

[(ii) の証明] $\mu(A)\leqq\alpha\mu(E)$ となる $A\subset E$ の全体を $\mathcal{A}$ とし, $\mathcal{A}$ の中に半順序 $\prec$ を

$$A_1\prec A_2 \iff \mu(A_1\setminus A_2)=0$$

でいれる. (ただし $\mathcal{A}$ の中の集合で μ 測度0を除いて一致するものは同一視する.) $\mathcal{A}'$ を $\mathcal{A}$ の中の全順序系とするとき, $\mathcal{A}'$ のすべての元より大きい $\mathcal{A}$ の元 A が存在することをいう. $\beta=\sup\{\mu(A')\,|\,A'\in\mathcal{A}'\}$ とすると, $\mathcal{A}'$ の中に増大列

$$A_1{}'\subset A_2{}'\subset\cdots, \qquad \mu(A_n{}')\longrightarrow\beta$$

がある. $A=\bigcup_n A_n{}'$ とおけば,

$$\mu(A)=\lim_n\mu(A_n{}')=\beta\leqq\alpha\mu(E)$$

となるから, $A\in\mathcal{A}$ となる. $\mathcal{A}'$ の中の任意の A' をとれば, 'ある n で $A'\prec A_n$', または 'すべての n に対し $A_n\prec A'$' のいずれかである. 前者のばあいは当然 $A'\prec A$ である. 後者のばあいには, $A\prec A'$ となり $\mu(A)=\beta\geqq\mu(A')$ により, $\mu(A'\setminus A)=0$ すなわち $A'\prec A$ となる. かくして $\mathcal{A}$ の任意の全順序部分系 $\mathcal{A}'$ は $\mathcal{A}$ のある元 A でおさえられることがわかったから, Zorn の補題により, $\mathcal{A}$ に極大元がある. $\mathcal{A}$ の極大元の一つを E' とするとき, $\mu(E')\leqq\alpha\mu(E)$ は明らかであるが, 実は等式がなりたつ. もし, そうでなければ, $\mu(E-E')>(1-\alpha)\mu(E)>0$ により, (i) を用いて $E''\subset E-E'$ をもとめて

$$0<\mu(E'')<\alpha\mu(E)-\mu(E')$$

とすることができる. $E'+E''\subset E$ でしかも

$$\mu(E'+E'')=\mu(E')+\mu(E'')<\alpha\mu(E)$$

であるから, $E'+E''$ は $\mathcal{A}$ に属することになる. E' は $\mathcal{A}$ の極大元であるから,

$E' \succ E'+E''$, したがって $\mu(E'')=0$ となり, 矛盾となる. ゆえに $\mu(E')=\alpha\mu(E)$ でなければならない. ▌

[(iii) の証明] (ii) により $\mu(E)/n$ に等しい測度をもつ $E_1 \subset E$ がある. 再び (ii) を用いて, $\mu(E-E_1)/(n-1)=\mu(E)/n$ に等しい測度をもつ $E_2 \subset E-E_1$ がみいだされる. これをくりかえして, E の分割

$$E = \sum_{k=1}^{n} E_k, \qquad \mu(E_k) = \frac{\mu(E)}{n} \quad (k=1,2,\cdots,n)$$

を得る. n を十分大きくとれば, (iii) の分割が得られる. ▌

$(S, \mathcal{S})$ を Borel 空間とする. 非負整数値 (∞ を許す) をとる $(S, \mathcal{S})$ の上の測度 m を S の上の**配置**という. 配置 m は S の上の測度であるが, これから S の上の点関数 $m(\{x\})$ $(x \in S)$ が得られる. 集合

$$A = \{x \mid m(\{x\})>0\}$$

を配置 m の**台**という. m の σ 有限性により, $m(\{x\})$ は常に有限であり,

$$m(E) = \sum_{x \in A \cap E} m(\{x\}), \qquad E \in \mathcal{S}$$

となり, m は $m(\{x\})$ から再生される. とくにすべての $x \in A$ に対して $m(\{x\})=1, 0$ のとき, 配置 m を**固有配置**という.

駐車場の駐車状態は配置であらわされる. 1台の車に対する駐車区画の全体を S とし, $E \subset S$ の中の駐車台数を $m(E)$ であらわすと, m は S の上の配置である. $m(\{x\})$ は区画 x に駐車している台数をあらわし, m の台 A はふさがっている区画の全体である. m が固有配置となっているのは, 1区画に2台以上駐車 (違法駐車) しているものがないことを意味する.

$(S, \mathcal{S})$ を Borel 空間とし, $E \in \mathcal{S}$ に依存する確率変数系 $X=\{X(E,\omega), E \in \mathcal{S}\}$ が $(S, \mathcal{S})$ の上の**偶然配置**であるとは, ほとんど確実に $X(E,\omega)$ が E の関数として $(S, \mathcal{S})$ の上の配置となっていることである. このとき

$$\mu(E) = E(X(E)), \qquad E \in \mathcal{S}$$

は S の上の測度となるが, 必ずしも σ 有限ではない. μ を偶然配置 X の**平均測度**という. ここでは μ の Lebesgue 拡大 (これも μ とかく) が前述の意味で $(S, \mathcal{S})$ の上の測度であることを仮定する.

$(S, \mathcal{S})$ の上の偶然配置 $X=\{X(E,\omega), E \in \mathcal{S}\}$ が**多項配置**であるとは, S の任意の分割

$$S = \sum_{r=1}^{d} E_r \qquad (E_r \in \mathcal{S})$$

に対し, $(X(E_1), X(E_2), \cdots, X(E_d))$ の結合分布が d 次元の多項分布 (§2.8) となることであると定義する. μ を X の平均測度とする. 特別な分解 $S=S$ を考えて見ると, $X(S)$ は1次元の多項分布 (すなわち, ある自然数に集中した δ 分布) に従うから,

$$X(S) = E(X(S)) = \mu(S) \quad \text{a.s.}$$

でなければならない. これは平均測度 μ の全測度 $\mu(S)$ が自然数であることを意味する. したがって上の $X(E_1), X(E_2), \cdots, X(E_d)$ の結合分布は

$$P_{p_1, p_2, \cdots, p_d, k} \qquad (k=\mu(S))$$

の形をしている. これから

$$E(t_1^{X(E_1)} t_2^{X(E_2)} \cdots t_d^{X(E_d)}) = (p_1 t_1 + p_2 t_2 + \cdots + p_d t_d)^k$$

が得られる. 両辺を t_r で偏微分して $t_1=t_2=\cdots=t_d=1$ とおいて

$$EX(E_r) = kp_r \quad \text{すなわち} \quad p_r = \frac{\mu(E_r)}{\mu(S)}$$

を得る. 以上を綜合すると, 多項配置 X の平均測度 μ の全測度 $\mu(S)$ は自然数で, 任意の分割 $S=\sum_{r=1}^{d} E_r$ に対し $X(E_r)$, $r=1,2,\cdots,d$ の結合分布は多項分布

$$P_{p_1, p_2, \cdots, p_d, k} \qquad \left(p_r = \frac{\mu(E_r)}{\mu(S)}, \ k=\mu(S)\right)$$

である.

上述のことから多項配置は, 法則同等を除いて, 平均測度によって完全に定まる. さて $(S, \mathcal{S})$ 上の任意の測度 μ (ただし $\mu(S)$ は自然数) に対して, これを平均測度とするような多項配置が存在することを示そう.

$\nu(E)=\mu(E)/\mu(S)$ は (完全可分な) 確率測度であるから,

$$(\Omega, P) = (S^k, \nu^k) \qquad (k=\mu(S))$$

とおくと, 確率空間が得られ,

$$Y_i(\omega) = \omega \text{ の } i \text{ 座標}, \qquad i=1,2,\cdots,k$$

とおくと, これは独立な S 値確率変数である. しかもすべての Y_r は ν に従っている. さて

$$X(E,\omega)=\sum_{i=1}^{k}1_E(Y_i(\omega))$$

$$(=Y_1,\,Y_2,\cdots,Y_k\text{ の中で }E\text{ に入るものの数})$$

とおくと，$X=\{X(E,\omega),E\in\mathcal{B}\}$ は平均測度 μ の多項配置である．実際，分割 $S=\sum_{r=1}^{d}E_r$ に対し，

$$E\left(\prod_{r=1}^{d}t_r^{X_r}\right)=E\left(\prod_{r=1}^{d}\prod_{i=1}^{k}t_r^{I_{r,i}}\right)\qquad (X_r=X(E_r),\ \ I_{r,i}=1_{E_r}(Y_i))$$

$$=\prod_{i=1}^{k}E\left(\prod_{r=1}^{d}t_r^{I_{r,i}}\right)\qquad (\{Y_i\}\text{ の独立性により})$$

$$=\prod_{i=1}^{k}\left(\sum_{r=1}^{d}t_rP\{I_{r,i}=1\}\right)\qquad \left(\begin{matrix}\sum_{i=1}^{d}I_{r,i}=1\text{ により，}I_{r,i}=1\\ \text{ならば }I_{s,i}=0\ (s\neq r)\end{matrix}\right)$$

$$=\prod_{i=1}^{k}\left(\sum_{r=1}^{d}t_rP\{Y_i\in E_r\}\right)=\left(\sum_{r=1}^{d}t_r\frac{\mu(E_r)}{\mu(S)}\right)^k$$

となる．$P\{X(E)=0,E\in\mathcal{B}\}=1$ のときにも，便宜上 X を平均測度 $\mu\equiv0$ の多項配置とよぶことにしておく．

多項配置についてはこの辺でとどめて，つぎに Poisson 配置を定義しよう．$X=\{X(E,\omega),E\in\mathcal{B}\}$ を $(S,\mathcal{B})$ の上の偶然配置とし，μ をその平均測度とする．X がつぎの2条件をみたすとき，X を **Poisson 配置**という：

(i) 任意の $E\in\mathcal{B}$ に対し，$X(E)$ が平均 $\mu(E)$ の Poisson 分布に従う ($\mu(E)=0$ のときには $X(E)=0$ a.s.，$\mu(E)=\infty$ のときには $X(E)=\infty$ a.s. とする)，

(ii) 任意の $E_1,E_2,\cdots,E_n\in\mathcal{B}$ $(E_i\cap E_j=\phi,\ i\neq j)$ に対しては，$X(E_i)$，$i=1,2,\cdots,n$ が独立である．

μ は σ 有限であるから，可算分割

$$S=\sum_n S_n,\qquad 0<\mu(S_n)<\infty,\qquad S_n\in\mathcal{B}$$

が存在する．上の条件 (ii) により

$$X_n=\{X_n(E)\equiv X(E),E\in\mathcal{B}\cap S_n\}\quad(=X\text{ の }S_n\text{ への制限}),\qquad n=1,2,\cdots$$

は独立である．しかも

$$P\{X_n(S_n)=k\}=P\{X(S_n)=k\}=e^{-\mu(S_n)}\frac{\mu(S_n)^k}{k!},\qquad k=0,1,2,\cdots.$$

$k>0$ のときには，S_n の任意の分割 $S_n=\sum_{r=1}^{d}E_r$ $(E_r\in\mathcal{B})$ と任意の $k_1,k_2,\cdots,k_d$ $\left(\sum_r k_r=k\right)$ に対し

$$P\{X_n(E_r)=k_r,r=1,2,\cdots,d\,|\,X(S_n)=k\}$$

$$=\prod_{r=1}^{d}e^{-\mu(E_r)}\frac{\mu(E_r)^{k_r}}{k_r!}\Big/e^{-\mu(S_n)}\frac{\mu(S_n)^k}{k!}$$

$$=\frac{k!}{k_1!k_2!\cdots k_d!}\prod_{r=1}^{d}\left(\frac{\mu(E_r)}{\mu(S_n)}\right)^{k_r}\qquad\left(\because\ \mu(S_n)=\sum_r\mu(E_r)\right).$$

これは，X_n が $(\Omega,P_{X(S_n)=k})$ の上で平均測度 $k\mu(\cdot\cap S_n)/\mu(S_n)$ の多項配置となることを示している．

これだけのことを念頭において平均測度 μ の Poisson 配置を構成しよう．S を上のように分解し，まず固定した n に対し X の S_n における制限 X_n を構成しよう．平均 $\mu(S_n)$ の Poisson 分布に従う確率変数 Z と $\nu(E)=\mu(E)/\mu(S_n)$ $(E\in\mathcal{B}\cap S_n)$ に従う確率変数列 $Y_1,Y_2,\cdots$ をとり，しかも $Z,Y_1,Y_2,\cdots$ が独立となるようにする．

$$X_n(E)=\sum_{i=1}^{Z}1_E(Y_i),\qquad E\in\mathcal{B}\cap S_n$$

とおくと，これは $(S_n,\mathcal{B}\cap S_n)$ の上の偶然配置で，

$$P\{X_n(E_r)=k_r,r=1,2,\cdots,d\}$$

$$=P\{X_n(S_n)=k,X_n(E_r)=k_r,r=1,2,\cdots,d\}\qquad\left(k=\sum_{r=1}^{d}k_r\right)$$

$$=P\left\{Z=k,\sum_{i=1}^{k}1_{E_r}(Y_i)=k_r,r=1,2,\cdots,d\right\}$$

$$=P\{Z=k\}P\left\{\sum_{i=1}^{k}1_{E_r}(Y_i)=k_r,r=1,2,\cdots,d\right\}$$

$$=e^{-\mu(S_n)}\frac{\mu(S_n)^k}{k!}\frac{k!}{k_1!\cdots k_d!}\left(\frac{\mu(E_1)}{\mu(S_n)}\right)^{k_1}\cdots\left(\frac{\mu(E_d)}{\mu(S_n)}\right)^{k_d}$$

$$=\prod_{r=1}^{d}e^{-\mu(E_r)}\mu(E_r)^k/k_r!.$$

これから $d=1$ のときには

$$P\{X_n(E)=k\}=e^{-\mu(E)}\,\mu(E)^k/k_r!$$

がわかるから，これを上の式に入れて，

$$P\{X_n(E_r)=k_r,\ r=1,2,\cdots,d\} = \prod_{r=1}^{d} P\{X_n(E_r)=k_r\}$$

がでる．これで，おのおのの X_n が $(S_n, \mathcal{B}\cap S_n)$ の上の平均測度 $\mu|_{S_n}$ の Poisson 配置であることがわかった．$X_1, X_2, \cdots$ を独立となるように構成するには，常套の直積論法を用いる．さて

$$X(E,\omega) = \sum_n X_n(E\cap S_n, \omega)$$

とおけば，これは平均測度 μ の Poisson 配置となる．

平均測度 μ の Poisson 配置が二つ (X, X') あれば，任意の互いに素な $E_1, E_2, \cdots, E_d$ に対して $\{X(E_i)\}$, $\{X'(E_i)\}$ の結合分布は一致することは上述の論法を見れば，明らかであろう．$E_1, E_2, \cdots, E_d$ が互いに素でないときには，互いに素な $F_1, F_2, \cdots, F_m$ をとり，E_r を $\{F_i\}$ の中のものの和として表現すれば，$\{X(E_i)\}$, $\{X'(E_i)\}$ が同じ結合分布をもつことがわかる．

例題 5.9 $X=\{X(E), E\in\mathcal{B}\}$ を平均測度 μ の多項配置(または Poisson 配置)とする．$X(E)$ が E に関して固有配置である (確率が 1 である) ための必要十分条件は μ が連続であることを証明せよ．

[ヒント] 両配置の構成法から，つぎのことをいえば十分である．

Y_1, Y_2 が独立で S 上の同じ確率分布 ν に従うとき，

$$P\{Y_1=Y_2\}=0 \iff \nu \text{ の連続性.}$$

もし $\alpha\equiv\nu(\{a\})>0$ ならば

$$P\{Y_1=Y_2\} \geqq P\{Y_1=a,\ Y_2=a\} = \alpha^2 > 0.$$

もし ν が連続ならば，S の分解

$$S=\sum_{k=1}^{n} E_k, \qquad \nu(E_k)=\frac{1}{n} \quad (k=1,2,\cdots,n)$$

があることは，本節の冒頭に証明した．

$$P\{Y_1=Y_2\} \leqq P\left\{\bigcup_{k=1}^{n} \{Y_1\in E_k\}\cap\{Y_2\in E_k\}\right\}$$

$$\leqq \sum_{k=1}^{n} \nu(E_k)^2 = \frac{1}{n} \longrightarrow 0.$$

§5.10 加法過程

$X=\{X_t(\omega), t\in[0,\infty)\}$ を確率過程とする．$X_t(\omega)$ を $X_t, X(t), X(t,\omega)$ とかく

こともある．$X(0)\equiv 0$ でかつ任意の $0\leqq t_0<t_1<\cdots<t_n$ に対して，

$$X(t_k)-X(t_{k-1}), \quad k=1,2,\cdots,n \quad \text{が独立}$$

のとき，X を加法過程という．

例 5.1 $Y_0, Y_1, \cdots$ を独立な実確率変数列とし，

$$X_t(\omega)=\sum_{k\leqq t} Y_k(\omega)$$

とおくと，$X=\{X_t, t\in[0,\infty)\}$ は加法過程となる．

例 5.2 $Y_0, Y_1, \cdots$ を独立な実確率変数列とし，

$$\sum_{n=1}^{\infty}|EY_n|<\infty, \qquad \sum_{n=1}^{\infty}V(Y_n)<\infty$$

と仮定する．また $\Delta=\{s_1, s_2, \cdots\}$ を $[0,\infty)$ の中の任意に固定した可算列とする．

$$X_t(\omega)=\sum_{s_k\leqq t} Y_k(\omega)$$

とおく．上の和の意味は $s_k\leqq t$ をみたす k に対して $Y_k(\omega)$ を加えることを意味する．このとき $X=\{X_t, t\in[0,\infty)\}$ は加法過程となる．すべての t に対して上の和が概収束することは Kolmogorov の定理 4.3 により明らかである．

例 5.3 Wiener 過程は加法過程である．

例 5.4 $\{Y(A,\omega), A\in\mathcal{B}^1\cap[0,\infty)\}$ を $[0,\infty)$ の上の Poisson 配置とし，その平均測度が Lebesgue 測度に一致するとする．

$$X_t(\omega)=Y([0,t],\omega)$$

とおくと，$X=\{X_t(\omega), t\in[0,\infty)\}$ は加法過程となる．この加法過程は **Poisson 過程**とよばれる．ここで些細なことであるが，X_t の有限性について注意しておこう．上の定義からは，固定した n に対しては X_n は平均 n の Poisson 分布に従うから，$0\leqq X_n<\infty$ a.s. は明らか．したがって

$$0\leqq X_n<\infty, \qquad n=1,2,\cdots$$

がほとんど確実になりたつ．$X_t(\omega)$ は t について増大であるから，‘すべての t に対して $0\leqq X_t<\infty$’ がほとんど確実になりたつ．——

加法過程の中で特に重要なものは上の例 5.3, 例 5.4 のように，確率連続な加法過程である．以後 $X=\{X_t, t\in[0,\infty)\}$ は確率連続な加法過程であるとする．つぎの記号は今後しばしば有用である．

$$\sigma_{s,t}[dX]=\sigma[X_v-X_u,\ s\leqq u\leqq v\leqq t],$$

$$\sigma_{s+,t}[dX] = \sigma[X_v - X_u,\ s<u\leqq v\leqq t],$$
$$\sigma_{s,t-}[dX] = \sigma[X_v - X_u,\ s\leqq u\leqq v< t],$$
$$\sigma_{s+,t-}[dX] = \sigma[X_v - X_u,\ s<u\leqq v< t],$$
$$\mathcal{F}_{s,t}[dX] = \bigcap_{\varepsilon>0}\sigma_{s-\varepsilon,t+\varepsilon}[dX] \vee \mathbf{2}.$$

しばらくの間 X を固定して考えるから，$[dX]$ を省略する．

定理 5.25　X が確率連続な加法過程ならば，

(i)　$\mathcal{F}_{s,t} = \sigma_{s,t} \vee \mathbf{2}$.

(ii)　$s_0<s_1<\cdots<s_n$ ならば $\mathcal{F}_{s_{i-1},s_i}$, $i=1,2,\cdots,n$ は独立である．

証明　X の加法性を用いて σ_{s_{i-1},s_i}, $i=1,2,\cdots,n$ が独立であることは定理 3.11 から容易にわかる．したがって $\sigma_{s_{i-1},s_i}\vee \mathbf{2}$, $i=1,2,\cdots,n$ も独立である．これから (i) を証明すれば十分であることがわかる．

$$\mathcal{B}_t = \bigcap_{\varepsilon>0}\sigma_{t-\varepsilon,t+\varepsilon} \qquad (\text{ただし } \mathcal{B}_0=\bigcap_{\varepsilon>0}\sigma_{0,\varepsilon})$$

とおくと，

$$\mathcal{F}_{s,t} = \mathcal{B}_s \vee \sigma_{s,t} \vee \mathcal{B}_t \vee \mathbf{2}$$

となるから，(i) をいうには

$$\mathcal{B}_s \subset \mathbf{2}, \qquad s\in[0,\infty)$$

を証明すればよい．明らかに

$$\sigma_{s,t} = \sigma_{s,t-}\vee\sigma[X_t - X_s],$$
$$X_t - X_s = \operatorname*{l.i.p.}_{n\to\infty}\left(X\left(t-\frac{1}{n}\right)-X(s)\right)$$

であるから

$$\sigma_{s,t}\subset\sigma_{s,t-}\vee\mathbf{2}.$$

同様に

$$\sigma_{s,t}\subset\sigma_{s+,t}\vee\mathbf{2}.$$

さて $s>0$ のとき $\mathcal{B}_s\subset\mathbf{2}$ を証明しよう ($s=0$ のときの証明も同様で，むしろ簡単である)．s より大きい，t をとり，これを固定する．上に注意したように

$$\sigma_{0,s-\varepsilon},\qquad \sigma_{s-\varepsilon,s+\varepsilon},\qquad \sigma_{s+\varepsilon,t}$$

は独立である．$\mathcal{B}_s\subset\mathcal{B}_{s-\varepsilon,s+\varepsilon}$ により

$$\sigma_{0,s-\varepsilon},\qquad \mathcal{B}_s,\qquad \sigma_{s+\varepsilon,t}$$

も独立である．$\sigma_{0,s-\varepsilon}, \sigma_{s+\varepsilon,t}$ は $\varepsilon\downarrow 0$ のとき増大するから，定理 3.11 を用いて

$$\sigma_{0,s-},\qquad \mathcal{B}_s,\qquad \sigma_{s+,t}$$

の独立性がわかる．したがって

$$\sigma_{0,s}\ \ (\subset\sigma_{0,s-}\vee\mathbf{2}),\qquad \mathcal{B}_s,\qquad \sigma_{s,t}\ \ (\subset\sigma_{s+,t}\vee\mathbf{2})$$

が独立となる．ゆえに $\mathcal{B}_s$ は

$$\sigma_{0,s}\vee\sigma_{s,t}$$

と独立である．$\sigma_{0,s}\vee\sigma_{s,t}$ は定義により $\sigma_{0,t}$ と一致するから，$\mathcal{B}_s$ は $\sigma_{0,t}$ と独立である．$\mathcal{B}_s\subset\sigma_{0,t}$ であるから，$A\in\mathcal{B}_s$ のときには $A\in\sigma_{0,t}$ で，$\mathcal{B}_s$ と $\sigma_{s,t}$ の独立性により

$$P(A)=P(A\cap A)=P(A)P(A), \qquad \text{ゆえに}\quad P(A)=0 \ \text{または}\ 1$$

となる．これは $\mathcal{B}_s\subset\mathbf{2}$ を意味する．▌

$F[X]=\{\mathcal{F}_t[X]\}$ を X で生成される増大情報系とする．$X(0)\equiv 0$ であるから，

$$\mathcal{F}_t[X] = \mathcal{F}_{0,t}[dX]$$

となる．したがって上の定理により

$$\text{‘}s<t \Longrightarrow X_t - X_s \text{ は } \mathcal{F}_s[X] \text{ と独立’}$$

が得られる．したがって，複素数値をとる確率過程

$$Y_t^{\alpha} = \frac{e^{i\alpha X_t}}{E(e^{i\alpha X_t})}$$

は $F[X]=\{\mathcal{F}_t[X]\}$ に関してマルチンゲールとなる．(この意味は Y_t^{α} の実部，虚部がそれぞれマルチンゲールとなることである.) なぜならば，条件付確率も実部，虚部に分けて考えて，$t>s$ のとき

$$E(Y_t^{\alpha}\,|\,\mathcal{F}_s) = \frac{E(e^{i\alpha X_t}\,|\,\mathcal{F}_s)}{E(e^{i\alpha X_t})} \qquad (\mathcal{F}_s=\mathcal{F}_s[X]),$$
$$\begin{aligned} E(e^{i\alpha X_t}\,|\,\mathcal{F}_s) &= E(e^{i\alpha(X_t-X_s)}e^{i\alpha X_s}\,|\,\mathcal{F}_s)\\ &= E(e^{i\alpha(X_t-X_s)}\,|\,\mathcal{F}_s)e^{i\alpha X_s}\\ &= E(e^{i\alpha(X_t-X_s)})e^{i\alpha X_s} \qquad (\because\ X_t-X_s \text{ が } \mathcal{F}_s \text{ と独立}),\end{aligned}$$
$$E(e^{i\alpha X_t}) = E(e^{i\alpha(X_t-X_s)})E(e^{i\alpha X_s}) \qquad (\because\ X_t-X_s \text{ が } X_s \text{ と独立}),$$

したがって

$$E(Y_t^{\alpha}\,|\,\mathcal{F}_s) = e^{i\alpha X_s}/E(e^{i\alpha X_s}) = Y_s^{\alpha}$$

が得られ，Y_t^{α} のマルチンゲール性がわかる．

$Y^a=\{Y_t^a, t\in[0,\infty)\}$ の実部，虚部の D 変形(定理 5.19 註)をとって，Y^a と同等な複素数値 D 過程(実部，虚部が D 過程であるような複素数値過程) Z^a が存在する．X が確率連続であるから，$E(e^{iaX_t})$ は t について連続で，$\{E(e^{iaX_t})Z_t^a, t\in[0,\infty)\}$ も D 過程となる．これは $\{e^{iaX_t}, t\in[0,\infty)\}$ の D 変形である．したがってその虚部 $\{\sin aX_t, t\in[0,\infty)\}$ も D 変形がある．これを $\{U_t^a, t\in[0,\infty)\}$ としよう．

$X=\{X_t, t\in[0,\infty)\}$ が確率連続な加法過程でかつ D 過程でもあるとき，**Lévy 過程**という．Lévy 過程については，その構造が，あとで示すように，完全に決定されている．しかし，つぎの定理によって確率連続な加法過程の性質は Lévy 過程の性質に帰着される．

定理 5.26 確率連続な加法過程 $X=\{X_t, t\in[0,\infty)\}$ は，(強同等を除いて一意に定まる) Lévy 過程と同等である．

証明 X が D 変形をもつことをいえば，あとは容易である．まず r は有理数だけを動くとし

(i) $M_n\equiv\sup_{0\leqq r\leqq n}|X_r|<\infty$ a.s.

を証明しよう．そのためには，任意の $t\in[0,\infty)$ に対し，t', t'' ($t'<t<t''$，ただし $t=0$ のときには，$t'=0$ とする)が存在して

(ii) $\sup_{t'\leqq r\leqq t''}|X_r|<\infty$ a.s.

をいえば，Borel の被覆定理を用いて，上の '$M_n<\infty$ a.s.' がでる．

X は確率連続であるから，任意の a に対し t', t'' ($t'<t<t''$，ただし $t=0$ のときには $t'=0$) が存在して

$$P\{|X(t'')-X(s)|\leqq a\}>\frac{1}{2},\qquad s\in[t', t'']$$

がなりたつ．$\{r_n\}$ を $[t', t'']$ の中の稠密点列と，$t', r_1, r_2, \cdots, r_{n-1}, t''$ を大きさの順にならべて

$$t'=s_0\leqq s_1\leqq\cdots\leqq s_{n+1}=t''$$

とする．X の加法性により

$$X_k=X(s_k)-X(s_{k-1}),\qquad k=1,2,\cdots,n$$

は独立であり

$$X(s_k)-X(t')=X_1+\cdots+X_k,\qquad X(t'')-X(s_k)=X_{k+1}+\cdots+X_n$$

であるから，Ottaviani の不等式を用いて，$a\geqq1$ のとき

$$P\left\{\max_{1\leqq k\leqq n+1}|X(s_k)-X(t')|>2a\right\}\leqq 2P\{|X(t'')-X(t')|>a\}$$

がなりたつ．したがって

$$P\left\{\max_{1\leqq k\leqq n}|X(r_k)-X(t')|>2a\right\}\leqq 2P\{|X(t'')-X(t')|>a\}.$$

$n\to\infty$ として

$$P\left\{\sup_{t'\leqq r\leqq t''}|X(r)-X(t')|>2a\right\}\leqq 2P\{|X(t'')-X(t')|>a\}.$$

$a\to\infty$ として，

$$P\left\{\sup_{t'\leqq r\leqq t''}|X_r-X_{t'}|=\infty\right\}=0.$$

また $|X_{t'}|<\infty$ a.s. であるから，上の式から(ii)が得られ，さらに(i)が得られる．

$$\tau_n=\tau_n(\omega)=\inf\{r:|X_r|\geqq n\},\qquad n=1,2,\cdots$$

とおくと，(ii) により

$$\tau_1\leqq\tau_2\leqq\cdots,\quad \tau_n\longrightarrow\infty \text{ a.s.}$$

となる．P 零集合を除外して，これが例外なしになりたつと考えてよい．定理の前に導入した D 過程 U^a を考えると，

$$\sin aX_t=U_t^a \text{ a.s.},\qquad t\in[0,\infty)$$

であるから，P 零集合を除外して

$$\sin aX_r=U_r^a,\qquad r\in[0,\infty)\cap \boldsymbol{Q}$$

がすべての ω に対してなりたつとしてよい．この式から

$$r<\tau_n \Longrightarrow |X_r|<n \Longrightarrow \left|\frac{1}{n}X_r\right|<1<\frac{\pi}{2} \Longrightarrow X_r=n\arcsin U_r^{1/n}$$

が得られる．しかも $n<m$ のときには

$$r<\tau_n \Longrightarrow r<\tau_m \Longrightarrow m\arcsin U_r^{1/m}=X_r=n\arcsin U_r^{1/n}.$$

U^a が D 過程であるから，$m\arcsin U_t^{1/m}$，$n\arcsin U_t^{1/n}$ も D 過程で，上の関係から，$n<m$ のとき

$$t<\tau_n \Longrightarrow m\arcsin U_t^{1/m}=n\arcsin U_t^{1/n},$$

したがって，

$$\tilde{X}_t = n \arcsin U_t^{1/n} \quad (t<\tau_n), \qquad n = 1, 2, \cdots$$

とおくと，$\tau_n \to \infty$ により，$\tilde{X}_t$ はすべての t に対して，確定し，しかも

$$\tilde{X}_r = X_r, \qquad r \in [0, \infty) \cap \boldsymbol{Q}.$$

また $\tilde{X} = \{\tilde{X}_t, t \in [0, \infty)\}$ が確率連続な D 過程であることも明らかである．X は確率連続であるから，上の式から

$$\tilde{X}_t = X_t \text{ a.s.}$$

が得られ，$\tilde{X}$ が X の D 変形であることがわかった．▮

上の定理により確率連続な加法過程を調べるには，Lévy 過程を研究すればよいことがわかる．Lévy 過程の中で，見本過程がほとんど確実に連続なものを **Gauss 型の Lévy 過程**といい，また見本過程がほとんど確実に飛躍1で増加する階段関数のときには **Poisson 型の Lévy 過程**という．この名称はつぎの定理により正当化される．

定理 5.27　(i)　X が Lévy 過程で，かつ連続過程であれば，$X_b - X_a$ $(b>a)$ は Gauss 分布に従う．

(ii)　X が Lévy 過程で，しかもその見本過程がほとんど確実に飛躍1で増加する階段関数であれば，$X_b - X_a$ $(b>a)$ は Poisson 分布に従う．

証明　(i)　$X_t(\omega)$ はすべての ω に対して t に関して連続と仮定してよい．すべての ω に対し $X_t(\omega)$ は $[a, b]$ の上で一様連続であるから，

$$V_\varepsilon(\omega) = \sup \{|X_t(\omega) - X_s(\omega)| \mid t, s \in [a, b], |t-s| < \varepsilon\}$$

とおくと，すべての ω に対し

$$V_\varepsilon(\omega) \longrightarrow 0 \qquad (\varepsilon \to 0).$$

したがって $\varepsilon(n) \to 0$ を適当にとって

$$P\left\{V_{\varepsilon(n)}(\omega) > \frac{1}{n}\right\} \longrightarrow 0$$

とすることができる．n に対して $[a, b]$ の分点

$$a = a_0 < a_1 < \cdots < a_{m(n)} = b, \qquad a_k - a_{k-1} < \varepsilon_n \quad (k=1, 2, \cdots, m(n))$$

をとり，

$$|X(a_k) - X(a_{k-1})| \leqq \frac{1}{n} \quad \text{または} \quad > \frac{1}{n}$$

に応じて

$$X_{nk} = X(a_k) - X(a_{k-1}) \quad \text{または} \quad 0$$

とおき

$$S_n = \sum_{k=1}^{m(n)} X_{nk}$$

とおくと，$V_{\varepsilon(n)} \leqq 1/n$ のときには $S_n = X(b) - X(a)$ となるから，

$$P\{S_n \neq X(b) - X(a)\} \longrightarrow 0 \quad (n \to \infty), \qquad \text{ゆえに} \quad S_n \longrightarrow X(b) - X(a) \text{ i.p.}$$

したがって $\{S_n\}$ またはその部分列の確率法則が Gauss 分布に収束することをいえば，(i) の証明が完了する．

$V(S_n) \to \infty$ ならば，$X_{nk}' = (X_{nk} - E(X_{nk}))/V(S_n)^{1/2}$，$k=1, 2, \cdots, m(n)$ に定理 4.17 を適用して $(S_n - E(S_n))/V(S_n)^{1/2}$，$n=1, 2, \cdots$ の確率法則が $N_{0,1}$ に収束する．S_n は確率収束，$V(S_n) \to \infty$ により，$S_n/V(S_n)^{1/2}$ は 0 に確率収束，したがって $E(S_n)/V(S_n)^{1/2}$，$n=1, 2, \cdots$ の確率法則が $N_{0,1}$ に収束する．これは矛盾である．ゆえに $V(S_n)$ は有界な部分列をもち，当然収束部分列をもつ．$\{S_n\}$ の部分列をとることにより，$V(S_n) \to v \in [0, \infty)$ としてよい．

もし $|E(S_n)| \to \infty$ ならば，すべての $c>0$ に対し

$$P\{|S_n| < c\} \leqq P\{|S_n - E(S_n)| > |E(S_n)| - c\}$$
$$\leqq \frac{V(S_n)}{(|E(S_n)| - c)^2} \longrightarrow 0 \qquad (n \to \infty).$$

$\{S_n\}$ は $X(b) - X(a)$ に確率収束するから，

$$P\{|X(b) - X(a)| < c\} = 0, \qquad c > 0$$

となり，$c \to \infty$ として，矛盾に到達する．ゆえに $\{E(S_n)\}$ も有界部分列，したがって当然収束部分列をもつ．$\{S_n\}$ の部分列をとることにより，$E(S_n) \to m \in (-\infty, \infty)$ としてもよい．

$E(S_n) \to m$，$V(S_n) \to v$，$\max_k |X_{nk}| \leqq 1/n \to 0$ により，定理 4.17 を用いて，S_n の確率法則が $N_{m,v}$ に近づくことがわかる．ゆえに $X(b) - X(a)$ の確率法則は $N_{m,v}$ に等しい．

(ii) の証明も上と同様の方法で証明される．ただし中心極限定理のかわりに Poisson 少数法則を使う．▮

一般の Lévy 過程の構造を完全に決定しようというのが本節の主目的であるが，

その前に少し準備をしよう．$[0,\infty)$ の中の区間 $[s,t)$ に対応して $\mathcal{D}(P)$ の部分 σ 加法族 $\mathcal{B}_{s,t}$ が与えられているとする．$\{\mathcal{B}_{s,t}, 0\leqq s<t<\infty\}$ が加法系であるとは，

(A. 1) $\qquad s<t<u \Longrightarrow \mathcal{B}_{s,u}=\mathcal{B}_{s,t}\vee\mathcal{B}_{t,u}$,

(A. 2) $\quad [s_i,t_i),\ i=1,2,\cdots,n$ が互いに素 $\Longrightarrow \{\mathcal{B}_{s_i,t_i}, i=1,2,\cdots,n\}$ は独立

の2条件がなりたつことである．例えば $X=\{X_t\}$ が加法過程のとき，前に導入した

$$\sigma_{s,t}[dX]=\sigma[X_v-X_u, s\leqq u\leqq v\leqq t], \qquad 0\leqq s<t<\infty$$

は加法系である．これを'X で生成される加法系'という．

$\{\mathcal{B}_{s,t}\}$ を加法系とし，$X=\{X_t\}$ を任意の確率過程で，$X_0\equiv 0$ とする．もし

$$\sigma[X_t-X_s]\subset\mathcal{B}_{s,t}, \qquad 0\leqq s<t<\infty$$

ならば，当然 X は加法過程であって，しかも

$$\sigma_{s,t}[dX]\subset\mathcal{B}_{s,t}, \quad 0\leqq s<t<\infty$$

となる．このとき X は $\{\mathcal{B}_{s,t}\}$ に従属する加法過程という．

Lévy 過程 $X=\{X_t\}$ の見本関数は第1種不連続であるから，その不連続点 α において $X_\alpha(\omega)-X_{\alpha-}(\omega)$ が定まる．第1種不連続性により絶対値がある正数 a より大きい飛躍は有限時区間に有限個しかないから，それを加えていくと，階段型の関数になる．$a\to 0$ とすれば，極限として飛躍だけで変化する関数が得られ，残りは連続的に変化するように思われるが，事情はそれ程簡単ではない．$a\to 0$ としたとき，飛躍が無限に多くなっていくので，簡単には極限が考えられないからである．この様子を詳しく調べるために，補題をいくつか証明しておく．

補題 5.1 $X=\{X_t\}$ が加法系 $\{\mathcal{B}_{s,t}\}$ に従属するとき，

$$X_\alpha(\omega)-X_{\alpha-}(\omega)\in E \qquad (E\in\mathcal{B}(\boldsymbol{R}^1-\{0\}))$$

となるような時点 $\alpha\in(s,t]$ の数 $N((s,t]\times E,\omega)$ は可測 $\mathcal{B}_{s,t}$ である．

証明 始めに $u>0$ に対し $N_u(s,t]=N((s,t]\times(u,\infty),\omega)$ が可測 $\mathcal{B}_{s,t}$ であることをいう．見本関数の第1種不連続性により，すべての $s<t$ によりつぎの式がなりたつ．

$$\{N_u(s,t)\geqq 1\}=\bigcup_m\bigcap_n\bigcup_{r,r'}\left\{X(r')-X(r)>u+\frac{1}{m}\right\}\in\mathcal{B}_{s,t},$$

ここで最後の和は

$$r',r\in\boldsymbol{Q}\cup\{t\}, \qquad s<r<r'\leqq t, \qquad r'-r<1/n$$

をみたすすべての r,r' に対する和である．

$$\{N_u(s,t)\geqq n+1\}=\bigcup_{s<r<t}[\{N_u(s,r)\geqq n\}\cup\{N_u(r,t)\geqq 1\}]$$

により，すべての $s<t$ に対し $\{N_u(s,t)\geqq n\}\in\mathcal{B}_{s,t}$ とすれば，$\mathcal{B}_{s,r}\vee\mathcal{B}_{r,t}=\mathcal{B}_{s,t}$ $(s<r<t)$ により

$$\{N_u(s,t)\geqq n+1\}\in\mathcal{B}_{s,t}$$

が得られる．これから

$$\{N_u(s,t)\geqq n\}\in\mathcal{B}_{s,t}, \qquad n=1,2,\cdots$$

がなりたち，$N_u(s,t)$ は可測 $\mathcal{B}_{s,t}$ である．したがって

$$N((s,t]\times(u,v])=N_u(s,t)-N_v(s,t)$$

も可測 $\mathcal{B}_{s,t}$ である．$N((s,t]\times E)$ が E に関して σ 加法的であることに留意し，Dynkin 族定理を用いて，上の補題が $E\subset\mathcal{B}[(0,\infty)]$ に対してなりたつ．$E\in\mathcal{B}[(-\infty,0)]$ に対しても，したがって $E\in\mathcal{B}(\boldsymbol{R}-\{0\})$ に対しても同様である．∎

補題 5.2 $X=\{X_t\}$ は Lévy 過程，$Y=\{Y_t\}$ は Poisson 型の Lévy 過程で，ともに加法系 $\{\mathcal{B}_{s,t}\}$ に従属すると仮定する．もし X, Y の見本関数がすべての ω に対し共通の飛躍時点をもたないならば，X, Y は独立である．

証明 まず $s<t$ を固定して

$$A=X_t-X_s, \qquad B=Y_t-Y_s$$

が独立であることを示す．

$$s=t_{n,0}<t_{n,1}<\cdots<t_{n,n}=t$$

を区間 $[s,t]$ の n 等分点とし

$$A_{nk}=X(t_{n,k})-X(t_{n,k-1}), \qquad B_{nk}=Y(t_{n,k})-Y(t_{n,k-1}),$$

$$A_{nk}'=1_0(B_{nk})A_{nk}, \qquad A_n'=\sum_{k=1}^n A_{nk}'$$

とおく．共通飛躍時点がないという仮定により，すべての ω に対し

$$A_n'\longrightarrow A.$$

n 個の非負整数の組 $\varepsilon=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$ に対し

$$\varphi_n(\varepsilon)=E(e^{izA_n'}, B_{n1}=\varepsilon_1, B_{n2}=\varepsilon_2,\cdots,B_{nn}=\varepsilon_n)P\{B=0\},$$

$$\psi_n(\varepsilon)=E(e^{izA}, B=0)\prod_k P\{B_{nk}=\varepsilon_k\}$$

を比較しよう．

従属性の仮定により，A_{nk}', B_{nk} は $\mathcal{B}(t_{n,k-1}, t_{n,k})$ に関して可測であるから，$\{\mathcal{B}_{s,t}\}$ の加法性により

$$(A_{nk}', B_{nk}), \qquad k=1,2,\cdots,n$$

は独立である．固定した $\varepsilon=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ に対し $\varepsilon_k=0$ をみたす k に対する積，和を $\prod^0, \sum^0$ であらわし，$\varepsilon_k>0$ をみたす k に対する積，和を $\prod^+, \sum^+$ であらわす．明らかに

$$\prod_k f_k(\varepsilon_k) = \prod_k{}^0 f_k(0) \prod_k{}^+ f_k(\varepsilon_k).$$

$B_{nk}\neq 0$, $=0$ に応じて $A_{nk}'=0$, A_{nk} であり，$B=0 \Leftrightarrow B_{n1}=B_{n2}=\cdots=B_{nn}=0$ であるから，

$$\begin{aligned}\varphi_n(\varepsilon) &= \prod_{k=1}^n E(e^{izA_{nk}'}, B_{nk}=\varepsilon_k) \prod_{k=1}^n P\{B_{nk}=0\} \\ &= \prod_k{}^0 E(e^{izA_{nk}}, B_{nk}=0) \prod_k{}^+ P\{B_{nk}=\varepsilon_k\} \prod_k{}^0 P\{B_{nk}=0\} \prod_k{}^+ P\{B_{nk}=0\}.\end{aligned}$$

第1因子と第4因子，第2因子と第3因子をまとめ，ψ_n についても同様にして

$$\varphi_n(\varepsilon) = \prod_k E(e^{izA_{nk}1_0(\varepsilon_k)}, B_{nk}=0) \prod_k P\{B_{nk}=\varepsilon_k\},$$

$$\psi_n(\varepsilon) = \prod_k E(e^{izA_{nk}}, B_{nk}=0) \prod_k P\{B_{nk}=\varepsilon_k\}.$$

$|\varepsilon|=\sum\varepsilon_k$ とおくと，$\varphi_n(\varepsilon)$ と $\psi_n(\varepsilon)$ の差違は，第1因子の積で $\varepsilon_k>0$ をみたす k に対するもので，そのような k の個数は $|\varepsilon|$ をこえることはできない．しかも $\varepsilon_k>0$ のときには

$$\begin{aligned}&|E(e^{izA_{nk}1_0(\varepsilon_k)}, B_{nk}=0) - E(e^{izA_{nk}}, B_{nk}=0)| \\ &\leqq E|1-e^{izA_{nk}}| \\ &\leqq \sup_{\substack{\alpha,\beta\in[s,t] \\ |\alpha-\beta|\leqq(t-s)/n}} E|1-e^{iz(X_\alpha-X_\beta)}| \qquad (\equiv\delta_n \text{ とおく}).\end{aligned}$$

$|e^{iz\theta}-1|\leqq|z\theta|\wedge 2\leqq(2+|z|)(|\theta|\wedge 1)$ によって，上の δ_n は

$$\sup_{\substack{\alpha,\beta\in[s,t] \\ |\alpha-\beta|\leqq(t-s)/n}} (2+|z|)\|X_\alpha-X_\beta\|_0$$

でおさえられる．X の確率連続性により，$\alpha\mapsto X_\alpha$ は $\|\ \|_0$ に関して連続，したがって $\alpha\in[s,t]$ で一様連続であるから，$\delta_n\to 0$ $(n\to\infty)$ がなりたつ．

$$|a_k|, |b_k| \leqq 1, \quad k=1,2,\cdots,n \Longrightarrow \left|\prod_k a_k - \prod_k b_k\right| \leqq \sum_k |a_k-b_k|$$

により

$$|\varphi_n(\varepsilon)-\psi_n(\varepsilon)| \leqq |\varepsilon|\delta_n \prod_k P\{B_{nk}=\varepsilon_{nk}\} = |\varepsilon|\delta_n P\{B_{n1}=\varepsilon_1, B_{n2}=\varepsilon_2, \cdots, B_{nn}=\varepsilon_n\}.$$

$\varphi_n(\varepsilon)$, $\psi_n(\varepsilon)$ の定義により

$$E(e^{izA_n'}, B=p)P\{B=0\} = \sum_{|\varepsilon|=p}\varphi_n(\varepsilon), \qquad E(e^{izA}, B=0)P\{B=p\} = \sum_{|\varepsilon|=p}\psi_n(\varepsilon).$$

したがって，上に得た不等式から

$$|E(e^{izA_n'}, B=p)P\{B=0\} - E(e^{izA}, B=0)P\{B=p\}| \leqq p\delta_n P\{B=p\},$$

$n\to\infty$ とすると $\delta_n\to 0$, $A_n'\to A$. したがって

$$E(e^{izA}, B=p)P\{B=0\} = E(e^{izA}, B=0)P\{B=p\}.$$

$p=0,1,2,\cdots$ に対して加えて

$$E(e^{izA})P\{B=0\} = E(e^{izA}, B=0).$$

この両式から

$$E(e^{izA}, B=p) = E(e^{izA})P\{B=p\}, \qquad p=0,1,2,\cdots$$

を得る．これから A, B の独立性が容易に示される．

さて $0=t_0<t_1<t_2<\cdots<t_n$ に対し

$$\mathcal{A}_k = \sigma[X(t_k)-X(t_{k-1})], \qquad \mathcal{B}_k = \sigma[Y(t_k)-Y(t_{k-1})], \qquad k=1,2,\cdots,n$$

とおくと，上に証明したことから，おのおのの k に対し $\mathcal{A}_k$, $\mathcal{B}_k$ は独立，しかも

$$\mathcal{C}_k \equiv \mathcal{A}_k \vee \mathcal{B}_k \subset \mathcal{B}(t_{k-1}, t_k)$$

により $\mathcal{C}_k$, $k=1,2,\cdots,n$ は独立．したがって

$$\mathcal{A}_1,\ \mathcal{B}_1,\ \mathcal{A}_2,\ \mathcal{B}_2,\ \cdots,\ \mathcal{A}_n,\ \mathcal{B}_n$$

が独立となる．これから

$$\mathcal{A} \equiv \bigvee_k \mathcal{A}_k, \qquad \mathcal{B} \equiv \bigvee_k \mathcal{B}_k$$

が独立となる．$\mathcal{A}=\sigma[X(t_k), k=1,2,\cdots,n]$, $\mathcal{B}=\sigma[Y(t_k), k=1,2,\cdots,n]$ であることは容易にわかるから，結局 X と Y とが独立となる．▌

補題 5.3　任意の $E\in\mathcal{B}(\boldsymbol{R}^1-(-a,a))$ (a は正定数)に対し，$N_E(t)\equiv N((0,t]\times E, \omega)$, $t\in[0,\infty)$ は $\{\mathcal{B}_{s,t}\}$ に従属する Poisson 型 Lévy 過程である．

証明　$\{N_E(t)\}$ の見本関数が飛躍1の第1種不連続関数であることは定義から明らか．任意の $\varepsilon\in(0,1)$ に対し

$$P\{|N_E(t)-N_E(t-)|>\varepsilon\} = P\{N_E(t)-N_E(t-)=1\}$$

$$= P\{X(t)-X(t-) \in E\} \leqq P\{|X(t)-X(t-)| \geqq a\}.$$

一般に D 過程 $\{Y(t)\}$ に対しては，t における確率連続性は '$P\{|Y(t)-Y(t-)|>\varepsilon\}=0$ $(\varepsilon>0)$' と同等であるから，上の式から N_E の確率連続性がでる．補題 5.1 により，N_E は $\{\mathcal{B}_{s,t}\}$ に従属する加法過程である．したがって N_E は Lévy 過程である．N_E の見本関数は飛躍 1 の階段関数であるから，N_E は Poisson 過程である．∎

さて，任意の $E \in \mathcal{B}(\boldsymbol{R}^1-(-a,a))$ $(a>0)$ に対し

$$S_E(t) = \sum_{s\leq t}(X(s)-X(s-))1_E(X(s)-X(s-)),$$

$$X_E(t) = X(t)-S_E(t)$$

とおく．S_E は大きさが E に属する X の飛躍のみを加えていくことにより得られる確率過程である．おのおの見本 ω に対し

$$S_E(t)-S_E(s) = \int_{(s,t]}(X(\alpha)-X(\alpha-))dN_E(\alpha)$$

$$= \lim_{\delta\downarrow 0}\int_{(s,t]}(X(\alpha)-X(\alpha-\delta))dN_E(\alpha)$$

$$= \lim_{\delta\downarrow 0}\lim_{n\to\infty}\sum_{k=1}^{n}(X(s_{n,k})-X(s_{n,k}-\delta))N_E((s_{n,k-1}, s_{n,k}])$$

（$\{s=s_{n,0}<s_{n,1}<\cdots<s_{n,n}=t\}$ は $(s,t]$ の n 等分点）

に注意し，補題 5.1 を用いて，S_E が $\{\mathcal{B}_{s,t}\equiv\mathcal{B}_{s,t}[dX]\}$ に従属する加法過程であることがわかる．補題 5.3 の証明で注意したことから，S_E は確率連続であるから，S_E は $\{\mathcal{B}_{s,t}\}$ に従属する Lévy 過程となる．したがって，$X_E=X-S_E$ も同様である．

補題 5.4 任意の互いに素な集合族 $E_1, E_2, \cdots, E_n \in \mathcal{B}(\boldsymbol{R}^1-(-a,a))$ $(a>0)$ に対し，$N_{E_1}, N_{E_2}, \cdots, N_{E_n}, X_E\left(E=\sum_{k=1}^{n}E_k\right)$ は独立である．

証明 補題 5.3 により N_{E_i} が $\{\mathcal{B}_{s,t}\}$ に従属する Poisson 過程である．X_{E_1} は上に注意したよう $\{\mathcal{B}_{s,t}\}$ に従属する Lévy 過程であり，定義から，N_{E_1} とは共通の飛躍時点をもたない．したがって補題 5.2 により，X_{E_1} と N_{E_1} とは独立である．ここで X のかわりに X_{E_1}，E_1 のかわりに E_2 を考えると，$X_{E_1+E_2}$ と N_{E_2} は独立で，$\{\mathcal{B}_{s,t}[dX_{E_1}]\}$ に従属することがわかる．($E_1\cap E_2=\phi$ に注意せよ．) X_{E_1} と N_{E_1} とは独立であるから，$(X_{E_1+E_2}, N_{E_2})$ は N_{E_1} と独立である．これから N_{E_1}, N_{E_2}, $X_{E_1+E_2}$ が独立であることがわかる．この論法をくりかえして，一般の n に対して補題が正しいことがわかる．∎

補題 5.5 上と同様な $\{E_i\}$ に対し $S_{E_1}, S_{E_2}, \cdots, S_{E_n}, X_E$ も独立である．

証明

$$S_{E_i}(t) = \int_{E_i}uN_{du}(t) = \lim_{m\to\infty}\sum_k \frac{k}{m}N_{E_i(m,k)}(t), \qquad E_i(m,k) = E_i\cap\left(\frac{k-1}{m}, \frac{k}{m}\right]$$

であり，各 m に対し $E_i(m,k)$，$i=1,2,\cdots$，$k=0,\pm1,\pm2,\cdots$ は互いに素でその和は E であるから，上の近似和 $S_{E_i}{}^m$，$i=1,2,\cdots,n$，および X_E は独立であることは前の補題からわかる．したがって $S_{E_1}, S_{E_2}, \cdots, S_{E_n}, X_E$ も独立となる．∎

以上の準備のもとに Lévy 過程 $X=\{X(t,\omega)\}$ の飛躍の構造を明らかにしよう．X の飛躍は飛躍時点 t と飛躍の大きさ $X(t,\omega)-X(t-,\omega)$ の組であらわされる．したがって，飛躍の全体 $J=J(\omega)$ は

$$J = J(\omega) = \{(t, X(t,\omega)-X(t-,\omega)) \mid X(t,\omega)-X(t-,\omega)\neq 0\}$$

で与えられる．これは $\Gamma=[0,\infty)\times(\boldsymbol{R}-\{0\})$ の可算部分集合である．これは見本 ω に関係するから，偶然集合 (random set) とよんでもよい．$B\in\mathcal{B}(\Gamma)$ に対して B の中に入る飛躍の数 $\#(J(\omega)\cap B)$ を $N(B,\omega)$ であらわそう．ω を固定したとき，$N(B,\omega)$ は B に関して非負整数値（∞ を許す）をとる測度となっていることも明らかである．とくに

$$B=(0,t]\times E, \qquad E\in\mathcal{B}(\boldsymbol{R}^1-(-a,a)) \quad (a>0)$$

のときには，$N(B,\omega)$ は補題 5.3 の $N_t(E,\omega)$ と一致する．実際 $N(B,\omega)$ は

$$N((0,t]\times E,\omega) = N_t(E,\omega), \qquad E\in\mathcal{B}(\boldsymbol{R}^1-(-a,a)) \quad (a>0)$$

によって定まる測度といってもよい．（この条件で $N(B,\omega)$ が定まることは Dynkin 族定理により明らかである．）$N_t(E,\omega)$ は ω に関して P 可測であるから，$N(B,\omega)$ も同様であることは，やはり Dynkin 族定理により明らかである．したがって

$$N = \{N(B,\omega), B\in\mathcal{B}(\Gamma)\}$$

は偶然配置 (§5.9) であり，しかも明らかに固有偶然配置である．

偶然配置 N の平均測度を n とする，すなわち

$$n(B) = E(N(B)).$$

明らかに

$$n(\{t\}\times(\boldsymbol{R}-\{0\})) = E\{N(\{t\}\times(\boldsymbol{R}-\{0\}))\}$$
$$= E\{1_{\boldsymbol{R}-\{0\}}(X(t)-X(t-))\} = P\{X(t)-X(t-)\neq 0\}$$
$$= 0$$

である．したがって n は連続測度である．$n(B)=\infty$ となることも可能である．

補題 5.6 $N=\{N(B,\omega)\}$ は n を平均測度とする Poisson 固有配置である．

証明 $N(B)$ が Poisson 分布(平均値は $0,\infty$ を許す)に従い，かつ互いに素な $B_1, B_2, \cdots, B_r$ に対して $N(B_1), N(B_2), \cdots, N(B_r)$ が独立であることを証明すればよい．$\mathcal{I}$ によって

$$(s,t]\times(u,v] \qquad (uv>0)$$

の形の集合の全体をあらわし，$\mathcal{A}$ によって $\mathcal{I}$ に属する集合の有限直和の全体をあらわす．

$$N((s,t]\times(u,v]) = N_{(u,v]}(t)-N_{(u,v]}(s)$$

であるから，補題 5.3 により $I\in\mathcal{I}$ に対しては $N(I)$ は Poisson 分布に従う．また $I_1, I_2, \cdots, I_r\in\mathcal{I}$ が互いに素であれば，細分することにより，

$$I_i = \sum_{j=1}^{m}(s_{ij},t_{ij}]\times(u_j,v_j] \qquad \left(=\sum_{j=1}^{m} I_{ij} \text{ とおく}\right),$$

$$(u_j,v_j], \quad j=1,2,\cdots,m \quad \text{は互いに素},$$

$$(s_{ij},t_{ij}], \quad i=1,2,\cdots,r \quad \text{は互いに素} \qquad (j=1,2,\cdots,m)$$

の形にかける．$s_{ij}=t_{ij}$ のときには $(s_{ij},t_{ij}]=\phi$ とすることはいうまでもなかろう．

$$N(I_{ij}) = N_{(u_j,v_j]}(t_{ij})-N_{(u_j,v_j]}(s_{ij})$$

により，補題 5.4 により，m 次元確率変数の系

$$(N(I_{1j}), N(I_{2j}), \cdots, N(I_{nj})), \qquad j=1,2,\cdots,m$$

は独立である．また補題 5.3 により，おのおのの j に対し，

$$N(I_{ij}), \qquad i=1,2,\cdots,r$$

も独立である．これから

$$N(I_{ij}), \qquad i=1,2,\cdots,r, \quad j=1,2,\cdots,m$$

が独立であることがわかる．したがって

$$N(I_i) = \sum_j N(I_{ij}), \qquad i=1,2,\cdots,r$$

が独立である．独立な Poisson 変数(Poisson 分布に従う確率変数)の和は Poisson 変数であるから，すべての $A\in\mathcal{A}$ に対し $N(A)$ は Poisson 変数である．また $A_1, A_2, \cdots, A_r\in\mathcal{A}$ が互いに素であれば，上の $\{I_i\}$ と同様の細分ができるので，$N(A_1), N(A_2), \cdots, N(A_r)$ が独立であることがわかる．

Poisson 変数の単調列の極限も Poisson 変数であり，しかも平均値の極限も極限変数と一致するから，任意の開集合 $G\subset\Gamma$，に対しても，$N(G)$ は平均値 $n(G)$ の Poisson 変数であることがわかる．したがって，コンパクト集合に対しても同様である．任意の $B\in\mathcal{B}(\Gamma)$，$n(B)<\infty$ に対しては開集合列 $\{G_n\}$ とコンパクト集合列 $\{K_m\}$ をとって

$$G_m\supset B\supset K_m, \qquad n(G_m)-n(K_m)\longrightarrow 0$$

とすると $N(G_m)\geqq N(B)\geqq N(K_m)$ で，しかも

$$E(N(G_m)-N(K_m)) = n(G_m)-n(K_m)\longrightarrow 0 \qquad (m\to\infty).$$

これから

$$N(B) = \lim_{m\to\infty} N(K_m) \quad (\text{増大極限}) \quad \text{a.s.}$$

が得られる．これから，$N(B)$ も平均 $n(B)$ の Poisson 変数であることがわかる．もし $n(B)=\infty$ であれば，コンパクト集合列

$$K_1\subset K_2\subset\cdots\subset B, \qquad n(K_m)\longrightarrow\infty$$

が存在する．これから

$$P\{N(B)<a\} \leqq P\{N(K_m)<a\}$$
$$= \sum_{r<a} e^{-n(K_m)}\frac{n(K_m)^r}{r!} < \frac{an(K_m)^a}{e^{n(K_m)}}$$
$$\longrightarrow 0 \qquad (m\to\infty)$$

が得られ，$N(B)=\infty$ a.s. となる．$n(B)=\lim n(K_m)=\infty$ であるから，$N(B)$ は平均 $n(B)$ の Poisson 変数である．これで補題の前半は証明された．

$B_1, B_2, \cdots, B_r\in\mathcal{B}(\Gamma)$ が互いに素であれば，B_i に対しコンパクト集合列

$$K_{i1}\subset K_{i2}\subset\cdots\subset B_i, \qquad n(K_{im})\longrightarrow n(B_i)$$

をとると，

$$N(B_i) = \lim_{m\to\infty} N(K_{im}), \qquad i=1,2,\cdots,r$$

となる．したがって補題の後半を示すには，コンパクト集合 $K_1, K_2, \cdots, K_r$ についていえばよい．$\{K_i\}$ は互いに素であるから，$A_{im}\in\mathcal{A}$，$i=1,2,\cdots,r$ をとって

$$A_{i1}\supset A_{i2}\supset\cdots\supset K_i, \qquad n(A_{im}-K_i)\longrightarrow 0 \quad (m\to\infty) \qquad (i=1,2,\cdots,r),$$

$$A_{1m}, A_{2m}, \cdots, A_{rm} \quad \text{は互いに素} \qquad (m=1, 2, \cdots)$$

とすることができるから，上に用いた論法により，

$$N(K_i) = \lim_{m\to\infty} N(A_{im}) \ \text{a.s.} \qquad (i=1, 2, \cdots, r).$$

$\{A_{1m}, A_{2m}, \cdots, A_{rm}\}$ は互いに素であるから，上に証明したように $\{N(A_{1m}), N(A_{2m}), \cdots, N(A_{rm})\}$ は独立，したがって $N(K_i)$, $i=1, 2, \cdots, r$ も独立である．▌

Poisson 配置 $N=\{N(B)\}$ を用いると，

$$S_E(t, \omega) = \int_{(0,t]}\int_E uN(dsdu), \qquad E \in \mathcal{B}(\boldsymbol{R}-(-a, a)) \quad (a>0)$$

とかけることは明らかである．この確率変数の特性関数は

$$\exp\left\{\int_{[0,t]}\int_E (e^{izu}-1)n(dsdu)\right\}$$

となる．これを示すには $S_E(t, \omega)$ が

$$\lim_{m\to\infty}\sum_m \frac{k}{m}N([0, t]\times E_{mk}), \qquad E_{mk} = E\cap\left(\frac{k-1}{m}, \frac{k}{m}\right]$$

に等しいこと，$N([0, t]\times E_{mk})$, $k=0, \pm1, \pm2, \cdots$ が独立な Poisson 変数であることに注意し，Poisson 分布の特性関数の式を思いおこせばよい．

補題 5.7　すべての $t\in[0, \infty)$ に対し

$$\int_{[0,t]}\int_{\boldsymbol{R}-\{0\}} (u^2\wedge 1)n(dsdu) < \infty.$$

証明　$n_t(E)=n((0, t]\times E)$ とおくと，上の式は

$$\int_{|u|>0} (u^2\wedge 1)n_t(du) < \infty$$

と同等である．これは2条件

$$\int_{1>|u|>0} u^2 n_t(du) < \infty, \qquad n_t(\boldsymbol{R}-(-1, 1)) < \infty$$

と同等である．$n_t(\boldsymbol{R}-(-1, 1))=\infty$ のときには，

$$N((0, t]\times(\boldsymbol{R}-(-1, 1))) = \infty \ \text{a.s.}$$

となるが，これは X の見本関数の第1種不連続性に反する．したがって第2の不等式は正しい．

第1の不等式を証明しよう．補題 5.5 により

$$S_\varepsilon(t) = S_{\boldsymbol{R}-(-\varepsilon,\varepsilon)}(t), \qquad X_\varepsilon(t) = X_{\boldsymbol{R}-(-\varepsilon,\varepsilon)}(t)$$

は独立で，$X(t)=S_\varepsilon(t)+X_\varepsilon(t)$ となる．したがって

$$E(e^{izS_\varepsilon(t)})E(e^{izX_\varepsilon(t)}) = E(e^{izX(t)}) \qquad (=\varphi(z) \text{ とおく}).$$

特性関数の絶対値は1以下であるから，

$$|\varphi(z)| \leqq |E(e^{izS_\varepsilon(t)})| = \left|\exp\int_{|u|>\varepsilon}(e^{izu}-1)n_t(du)\right|$$

$$= \exp\int_{|u|>\varepsilon}(\cos zu-1)n_t(du)$$

$$\leqq \exp\int_{1\geqq|u|>\varepsilon}(\cos zu-1)n_t(du) \qquad (\cos zu\leqq 1)$$

$$\leqq \exp\int_{1\geqq|u|>\varepsilon} -\frac{z^2u^2}{4}n_t(du) \qquad (|z|<1 \text{ のとき}).$$

もし

$$I \equiv \int_{1\geqq|u|>0} u^2 n_t(du) = \infty$$

ならば，上の最後の不等式で $\varepsilon\downarrow 0$ として，

$$|\varphi(z)| = 0 \qquad (0<|z|<1),$$

$z\to0$ として $\varphi(0)=0$ を得る．しかし $\varphi(z)$ も特性関数であるから，$\varphi(0)=1$ でなければならない．これは矛盾である．したがって $I<\infty$ でなければならない．▌

これで Lévy 過程 X の飛躍の構造が明らかになった．飛躍の大きさが絶対値において $1/n$ と1との間のものを加えあわせると，

$$S_n(t, \omega) = \int_{s\leqq t}\int_{1/n\leqq|u|<1} uN(dsdu, \omega) = S_E(t, \omega),$$

$$E = \left\{u: \frac{1}{n}\leqq|u|<1\right\}, \qquad n = 2, 3, \cdots$$

に等しい．すでに証明したようにこれは Lévy 過程である．したがって $n\geqq2$ に対し，

$$T_n(t, \omega) = S_n(t, \omega)-E(S_n(t, \omega))$$

$$= \int_{s\leqq t}\int_{1/n\leqq|u|<1} uN(dsdu, \omega) - \int_{s\leqq t}\int_{1/n\leqq|u|<1} un(dsdu)$$

は平均0の Lévy 過程である．同様に，$m>n$ のとき

$$T_m(t, \omega)-T_n(t, \omega) = S_E(t, \omega)-E(S_E(t, \omega)), \qquad E = \left(-\frac{1}{n}, \frac{1}{n}\right)-\left(-\frac{1}{m}, \frac{1}{m}\right)$$

も平均0の Lévy 過程である．

補題 5.8 $m>n$ のとき

$$P\left\{\sup_{s\le t}|T_m(s,\omega)-T_n(s,\omega)|>\varepsilon\right\}\leqq\frac{1}{\varepsilon^2}\int_{1/m\leqq|u|<1/n}u^2n_t(du).$$

ここに $n_t(E)=n((0,t]\times E)$ である．

証明 $T_m(s)-T_n(s)$ が D 過程であるから左辺は

$$\lim_{p\to\infty}P\left\{\max_{i=1}^{p}|T_m(r_i)-T_n(r_i)|>\varepsilon\right\}\qquad(\{r_i\}=[0,t]\cap\boldsymbol{Q})$$

に等しい．$0, r_1, r_2, \cdots, r_p, t$ を大きさの順にならべて $r_1', r_2', \cdots, r_p', r_{p+1}', r_{p+2}'$ とすると，$T_m(s)-T_n(s)$ の加法性と Kolmogorov の不等式により，

$$P\left\{\max_{i=1}^{p+2}|T_m(r_i')-T_n(r_i')|>\varepsilon\right\}\leqq\frac{1}{\varepsilon^2}E((T_n(t)-T_m(t))^2).$$

したがって

$$P\left\{\max_{i=1}^{p}|T_m(r_i)-T_n(r_i)|>\varepsilon\right\}$$
$$\leqq\frac{1}{\varepsilon^2}E((T_m(t)-T_n(t))^2)=\frac{1}{\varepsilon^2}V(S_m(t)-S_n(t)),$$

$$S_m(t)-S_n(t)=\int_{s\le t}\int_{1/m\le u<1/n}uN(dsdu)$$

に注意し，有限和で近似することにより，特性関数のときと同様に考えて

$$V(S_n(t)-S_m(t))=\int_{s\le t}\int_{1/m\leqq|u|<1/n}u^2n(dsdu)=\int_{1/m\leqq|u|<1/n}u^2n_t(du).$$

したがって

$$P\left\{\max_{i=1}^{p}|T_m(r_i)-T_n(r_i)|>\varepsilon\right\}\leqq\frac{1}{\varepsilon^2}\int_{1/m\leqq|u|<1/n}u^2n_t(du).$$

$p\to\infty$ として補題の不等式が得られる．▮

補題 5.7 により上の不等式の右辺は $n, m\to\infty$ のとき 0 に収束するから，$D[0, t]$ の中の値をとる確率変数列

$$\boldsymbol{T}_n(\omega)=(T_n(s,\omega), 0\leqq s\leqq t),\qquad n=2,3,\cdots$$

が D のノルム

$$\|\xi\|_\infty=\sup_{0\le s\le t}|\xi(t)|\qquad\left(\begin{array}{l}D[s,t]\text{ は }\|\ \|_\infty\text{ に関して非可分 Banach}\\ \text{空間となることに注意せよ}\end{array}\right)$$

に関して確率収束することがわかる．確率収束は実数の場合について定義したが，Banach 空間のときにも同様に定義される．$(D[0,t], \|\ \|_\infty)$ は非可分であるから，本講の定義では確率変数といえないが，ここでは P 可測関数を確率変数ということにする．

補題 5.9 すべての t に関して，ほとんど確実に，$T_n(s,\omega)$, $s\in[0,t]$, は $n\to\infty$ のとき一様収束する．

証明 補題の意味は，上述の用語によれば，$D[0,t]$ の中の値をとる確率変数列 $\boldsymbol{T}_n$, $n=2,3,\cdots$ が概収束することである．

$$T_{n+1}(t)-T_n(t)=S_{E_n}(t)-E(S_{E_n}(t)),\qquad E_n=\left(-\frac{1}{n},\frac{1}{n}\right)-\left(-\frac{1}{n+1},\frac{1}{n+1}\right)$$

に注意すれば，補題 5.5 から $\boldsymbol{T}_2, \boldsymbol{T}_3-\boldsymbol{T}_2, \boldsymbol{T}_4-\boldsymbol{T}_3, \cdots$ が独立なことがわかる．したがって P. Lévy の 3 収束同等定理における

確率収束 $\Longleftrightarrow$ 概収束

が，Banach 空間の値をとる確率変数に対してもなりたつことがいえれば，補題の証明が完成する．実数値のときの証明をよく見ると，Ottaviani の不等式が鍵となっている．この不等式は，絶対値をノルムでおきかえると，Banach 空間のときにもなりたつことはすぐにわかる．▮

さて $\phi(u)=u\wedge 1\ (u>0)$, $u\vee(-1)\ (u<0)$ とおいて

$$\int_{s\le t}\int_{|u|\ge 1/n}\phi(u)n(dsdu)\quad\text{すなわち}\quad\int_{|u|\ge 1/n}\phi(u)n_t(du)$$

が t に関して連続であることを示そう．もし $t=t_0$ で不連続ならば

$$n\left(\{t_0\}\times\left(-\frac{1}{n},\frac{1}{n}\right)^c\right)>0,$$

$$P\left\{N\left(\{t_0\}\times\left(-\frac{1}{n},\frac{1}{n}\right)^c\right)>0\right\}>0,$$

$$P\left\{|X(t_0)-X(t_0-)|>\frac{1}{n}\right\}>0$$

となるが，これは X の確率連続性に反する．

つぎに

$$U_n(t,\omega)=\int_{s\le t}\int_{|u|\ge 1/n}uN(dsdu)-\int_{s\le t}\int_{|u|\ge 1/n}\phi(u)n(dsdu)$$

とおく．第1項の見本関数は階段型で，その飛躍は $X(t,\omega)$ の見本関数の絶対値 $1/n$ 以上の飛躍と一致する．第2項は t の連続関数であるから，$U_n(t,\omega)$ の見本関数の飛躍についても同様のことがいえる．また

$$U_{n+1}-U_n = T_{n+1}-T_n, \qquad n=2,3,\cdots$$

に注意すれば，任意の t に対しほとんど確実に，$U_n(s,\omega)$ は $s\leqq t$ で一様収束する．このことから，ほとんど確実に $U_n(s,\omega)$ は s に関して広義一様収束する．この極限を $U_\infty(s,\omega)$ とすると，広義一様収束性により $U_\infty(s,\omega)$ の絶対値 $1/n$ 以上の飛躍は $U_n(s,\omega)$ のそれに一致し，したがって $X(s,\omega)$ のそれとも一致する（ほとんど確実に）．これから

$$G(t,\omega) = X(t,\omega)-U_\infty(t,\omega) = \lim_{n\to\infty}(X(t,\omega)-U_n(t,\omega))$$

は連続過程であることがわかる．

$$U_n(t,\omega) = S_{E_n}(t,\omega)-\int_{|u|\geqq 1/n}\phi(u)n_t(du), \qquad E_n = \left(-\frac{1}{n},\frac{1}{n}\right)^c$$

により，$X-U_n$ は $\mathcal{B}_{s,t}=\mathcal{B}_{s,t}[dX]$ に従属するから加法過程である．したがって $G=X-U_\infty$ も加法過程である．G は連続な加法過程であるから，定理5.27により Gauss 型の Lévy 過程であることがわかる．

定理 5.28（Lévy 過程の分解定理）　Lévy 過程 X に対して $\Gamma=[0,\infty)\times(\boldsymbol{R}-\{0\})$ の上の固有な Poisson 配置 N とこれと独立な Gauss 型 Lévy 過程 G が存在して

$$X(t,\omega) = G(t,\omega)+\lim_{n\to\infty}\left[\int_{s\leq t}\int_{|u|>1/n}(uN(dsdu,\omega)-\phi(u)n(dsdu))\right]$$

とかける．ここで

$$n(B) = E(N(B,\omega))$$

で，この n は

$$\int_{s\leq t}\int_{|u|>0}(u^2\wedge 1)n(dsdu) < \infty, \qquad 0\leqq t<\infty$$

をみたし，$n(\{t\}\times(\boldsymbol{R}^1-\{0\}))=0$ である．

証明　この定理の大部分はすでに証明されている．残っているのは N と G との独立性と

$$n(\{t\}\times(\boldsymbol{R}^1-\{0\})) = 0$$

である．すでに示した

$$n\left(\{t\}\times\left(-\frac{1}{n},\frac{1}{n}\right)^c\right) = 0$$

から $n\to\infty$ として後者が得られる．N と G との独立性を示そう．上に導入した Lévy 過程 U_∞ に対する飛躍の Poisson 配置は N と一致するから，$\sigma[N]\subset\sigma[U_\infty]$ であることは明らかである．U_∞ と G とは独立，したがって $\sigma[U_\infty]$ と $\sigma[G]$ とは独立であるから，$\sigma[N]$ と $\sigma[G]$ とは独立，すなわち N と G とは独立である．∎

Lévy の分解定理の連続部分 G は Gauss 型 Lévy 過程であるから

$$m(t) = E(G(t)), \qquad v(t) = V(G(t))$$

は有限確定する．$m(t)$ は連続関数，$v(t)$ は連続増大関数で，$m(0)=v(0)=0$ となることを証明しよう．最後の条件は明らかであり，G の加法性により $v(t)$ の増大性も容易にわかる．問題は連続性である．G は連続過程であるから $\varphi(t,z)\equiv E\{\exp(izG(t))\}$ は t に関して連続であり，

$$\varphi(t,z) = \exp\left\{im(t)z-\frac{v(t)}{2}z^2\right\}.$$

$$|\varphi(t,1)| = \exp\left\{-\frac{v(t)}{2}\right\}$$

を用いて，$v(t)$ の連続性がでる．ゆえに

$$\exp\{im(t)z\} = \varphi(t,z)\exp\left\{\frac{v(t)}{2}z^2\right\}$$

も t に関して連続となる．したがって

$$\exp\left\{-\frac{m(t)^2}{2}\right\} = \int_{-\infty}^{\infty}\exp\{im(t)z\}N_{0,1}(dz) \qquad (N_{0,1}\text{ は Gauss 分布})$$

により $m(t)^2$ も連続である．

$$\overline{m}(t) = \lim_{\epsilon\downarrow 0}\sup_{|s-t|<\epsilon}m(s), \qquad \underline{m}(t) = \lim_{\epsilon\downarrow 0}\inf_{|s-t|<\epsilon}m(s)$$

とおくと，$m(t)^2$ の連続性から

$$\overline{m}(t)^2 = \underline{m}(t)^2 = m(t)^2 < \infty.$$

$\exp\{im(t)z\}$ が t に関して連続であるから

$$(\overline{m}(t)-\underline{m}(t))z = 2\pi n(t,z) \qquad (n(t,z)\text{ は整数}).$$

t を固定して z を動かすと，左辺は連続的に動くから，整数 $n(t,z)$ は z に無関係でなければならない．したがって左辺も z に無関係となり，$\bar{m}(t)=\underline{m}(t)$ すなわち m の連続性がでる．

以上の考察を綜合すると，Lévy 過程 X に対しては飛躍の Poisson 配置 $N=N_X$ の平均測度 $n=n_X$，連続部分 $G=G_X$ の平均値関数 $m(t)=m_X(t)$，分散関数 $v(t)=v_X(t)$ が定まる．ここで n, m, v はつぎの3条件をみたす．

(a)　n は $\Gamma=(0,\infty)\times(\boldsymbol{R}-\{0\})$ の上の測度で，すべての t に対し，

$$n(\{t\}\times(\boldsymbol{R}-\{0\}))=0, \qquad \int_{s\leq t}\int_{|u|>0}(u^2\wedge 1)n(dsdu)<\infty,$$

(b)　$m(t)$ は連続，$m(0)=0$，

(c)　$v(t)$ は連続増大，$v(0)=0$.

この (n,m,v) を Lévy 過程 X の**特性量**という．逆にこの3条件をみたす (n,m,v) が与えられたとき，これを特性量とする Lévy 過程 X が存在するか，またただ1個しか存在しないかという問題がおこる．これがいえれば，特性量という名称が正当化されるわけである．

定理 5.29（**Lévy 過程の構成定理**）　上の3条件 (a), (b), (c) をみたす (n,m,v) が与えられたとき，$n=n_X$，$m=m_X$，$v=v_X$ となる Lévy 過程 X が存在し，しかもこの X は法則同等を除いて一意に定まる．

証明　一意性は分解定理からほとんど明らかである．実際このような X があれば X の飛躍の Poisson 配置 N は法則同等を除いてその平均測度 n によってきまる（§5.9）．また X の連続部分 G も法則同等を除いて (m,v) によりきまることは容易にわかる．しかも分解定理により N と G とは独立であるから，(N, G) も法則同等を除いて (n,m,v) によりきまる．分解定理により X は (N,G) であらわされるから，X も (n,m,v) により法則同等を除いてきまる．

存在を示すには，Wiener 過程 B と n を平均測度とする Poisson 配置 N を独立となるようにとり

$$X(t)=m(t)+B(v(t))+\lim_{n\to\infty}\int_{s\leq t}\int_{|u|>1/n}(uN(dsdu)-\phi(u)n(dsdu))$$

とおけばよい．n に関する条件 (a) は，N が‘すべての t に対し $N(\{t\}\times\boldsymbol{R}^1-\{0\})=0$’をほとんど確実にみたすことと上の積分がほとんど確実に t に関し広義一様

収束することを示すのに用いられ，m, v に関する条件 (b), (c) は $G(t)=m(t)+B(v(t))$ が連続（したがって Gauss 型）Lévy 過程となることを示すのに用いられる．詳細は，読者にまかせるが，Poisson 配置の性質と分解定理の証明をよく理解すれば，さほど困難ではなかろう．∎

例題 5.10　Lévy 過程に関するつぎの事実を証明せよ．

(i)　分解定理の $\phi(u)$ は $u/(1+u^2)$ でおきかえてもよい（これに応じて $m(t)$ は修正を受ける）ことを示せ．実際 Lévy はこれを用いている．もっと一般に‘$u=0$ の近傍で $\psi(u)=u+o(u^2)$’をみたす有界関数 $\psi(u)$ でおきかえてもよい．

［ヒント］　n に関する条件 (a) の後者（可積分条件）に注意せよ．

(ii)　Lévy 過程 X の特性量を (n,m,v) とすれば，$X(t)-X(s)$ $(t>s)$ の特性関数 $E[\exp\{iz(X(t)-X(s))\}]$ は

$$\exp\left\{i(m(t)-m(s))z-\frac{1}{2}(v(t)-v(s))z^2+\int_{|u|>0}(e^{izu}-1-i\phi(u)z)n((s,t]\times du)\right\}$$

に等しいことを示せ．

［ヒント］　分解定理の N による積分の部分を有限和で近似せよ．

(iii)　前問を用いて Lévy 過程が特性量により法則同等を除いて一意に定まること（構成定理の後半）を示せ．

［ヒント］　前問により $X(t)-X(s)$ の確率法則は (n,m,v) で定まる．X の加法性を用いて，有限時点 $\{t_i\}$ における $\{X(t_i)\}$ の結合分布も (n,m,v) で定まる．

(iv)　$0\leqq s\leqq t<\infty$ に依存する1次元分布の系 $\{\mu_{s,t}\}$ があり，$s\leqq t\leqq u$ のとき $\mu_{s,t}*\mu_{t,u}=\mu_{s,u}$ をみたすとする．$\mu_{s,t}$ が s,t について（弱位相に関し）連続とする．この $\{\mu_{s,t}\}$ に対し X_t-X_s $(t>s)$ の確率法則が $\mu_{s,t}$ に等しいような Lévy 過程 X が存在し，法則同等を除いて一意である．

［ヒント］　問題の条件をみたす確率連続な加法過程 X が存在するが，その D 変形（定理 5.26）をとれば求めるものとなる．一意性は X の加法性から明らかである．

(v)　X_t-X_s $(t>s)$ の確率法則が $t-s$ のみに関係するような Lévy 過程を（時間的に）**一様**な Lévy 過程という．このときには n_X, m_X, v_X はつぎの形にかけることを示せ．

$$m_X(t)=m\cdot t, \qquad v_X(t)=v\cdot t, \qquad n_X(dtdu)=dt\cdot n(du),$$

$$m \in R, \quad v \geq 0, \quad \int_{|u|>0} (u^2 \wedge 1) n(du) < \infty.$$

註 $n(du)$ を一様 Lévy 過程 X の Lévy 測度という.

［ヒント］ X の一様性，加法性により，$X^{(s)} = \{X(t+s) - X(s), 0 \leq t < \infty\}$ が X と法則同等であることに注意し

$$m_X(t+s) = m_X(t) + m_X(s), \qquad v_X(t+s) = v_X(t) + v_X(s),$$
$$n_X(t+s, E) = n_X(t, E) + n_X(s, E), \qquad n_X(t) = n_X((0, t] \times E)$$

を示し，m_X, v_X の連続性，$n_X(t, E)$ の非負性を用いて，これらが上の形にかけることを証明せよ.

§5.11 無限可分解分布

1次元の分布 μ に対し，$t \in [0, \infty)$ に依存する分布の系 $\{\mu_t\}$ をみいだして

$$\text{(a)} \quad \begin{cases} \mu_t \text{ は } t \text{ に関し（弱位相に関し）連続,} \\ \mu_t * \mu_s = \mu_{t+s}, \quad t, s \in [0, \infty), \\ \mu_0 = \delta \quad \text{（この条件は上の条件からでるから，除いてもよい）,} \\ \mu_1 = \mu \end{cases}$$

とすることができるとき，μ を**無限可分解**という．始めの3条件は $\{\mu_t\}$ がたたみこみに関して連続半群をなしていることを意味する.

無限可分解性はもっとゆるい条件で定義することもできる．例えば

(b) 任意の n に対して $\mu = \mu_n * \mu_n * \cdots * \mu_n$ (n 個) の形にかける.

(c) 任意の $\varepsilon > 0$ に対し，分解

$$\mu = \mu_1 * \mu_2 * \cdots * \mu_n,$$
$$d_L(\mu_i, \delta) < \varepsilon, \quad i = 1, 2, \cdots, n \quad (d_L \text{ は Lévy 距離})$$

が存在する.

(d) $\mu_n = \mu_{n1} * \cdots * \mu_{nm}$ $(m = m(n))$, $\max_k d_L(\mu_{nk}, \delta) \to 0$ が存在して $\mu_n \to \mu$.

これらの4条件は後のものほどゆるくなっていることは明らかであるが，実は同等である．同等性の証明はかなり複雑であるから省略し，一見最も強い条件で定義しておく.

§2.6 a) で示した

$$N_{m_1, v_1} * N_{m_2, v_2} = N_{m_1+m_2, v_1+v_2}, \qquad p_{\lambda_1} * p_{\lambda_2} = p_{\lambda_1+\lambda_2},$$

$$C_{m_1, c_1} * C_{m_2, c_2} = C_{m_1+m_2, c_1+c_2}$$

を用いると，Gauss 分布 $N_{m,v}$，Poisson 分布 p_λ，Cauchy 分布 $C_{m,c}$ が無限可解であることがわかる．例えば $\mu = N_{m,v}$ に対しては $\mu_t = N_{mt, vt}$ とおけば，条件(a)がなりたつ．μ が無限可解のとき，条件(a)にあらわれる μ_t はすべて無限可解であることは，半群 $\{\mu_{t,s}, s \in [0, \infty)\}$ を考えてみれば明らかであろう.

X が一様な Lévy 過程のときには，$X(t) - X(s)$ $(t \geq s)$ の確率法則は $t-s$ のみに依存するから，これを μ_{t-s} とかくと，X の確率連続性から μ_t は t に関して連続であり，X の加法性から $\mu_t * \mu_s = \mu_{t+s}$ がでるから，$\{\mu_t\}$ は連続半群となり，すべての t に対し，μ_t は無限可解である.

逆に分布の連続半群 $\{\mu_t\}$ が与えられているとする．当然 μ_t はすべて無限可解である．この半群に対し

$$\mu_{t,s} = \mu_{t-s}, \qquad s \leq t$$

とおくと，$\mu_{t,s}$ は t, s に関し連続で，

$$\mu_{s,t} * \mu_{t,u} = \mu_{s,u}, \qquad s \leq t \leq u$$

がなりたつから，前節の例題 5.10 (iv) により，Lévy 過程 X を構成して $X(t) - X(s)$ の確率法則を $\mu_{s,t}$ に等しくすることができる．$\mu_{s,t} = \mu_{t-s}$ であるから，X は一様な Lévy 過程となり，$X(t) - X(s)$ は μ_{t-s} に従う．X の特性量 $m_X(t)$, $v_X(t)$, $n_X(dtdu)$ はつぎの形で与えられる（前節例題 5.10 (v)).

$$m_X(t) = m \cdot t, \qquad v_X(t) = v \cdot t, \qquad n_X(dtdu) = dtn(du).$$

ここで

$$m \in R, \quad v \geq 0, \quad \int_{|u|>0} (u^2 \wedge 1) n(du) < \infty.$$

これから，μ_t の特性関数 $\mathscr{F}\mu_t$ はつぎの形にあらわされることがわかる（前節例題 5.10 (ii)).

$$\mathscr{F}\mu_t(z) = E[e^{izX(t)}] = \exp(t\psi(z)).$$

ここで

$$\psi(z) = imz - \frac{v}{2} z^2 + \int_{|u|>0} (e^{izu} - 1 - i\phi(u) z) n(du).$$

定理 5.30（無限可解分布の Lévy の分解定理） 任意の無限可解分布 μ はつぎの式で定められる.

$$\mathcal{F}\mu(z) = e^{\psi(z)}.$$

ここで

$$\psi(z) = imz - \frac{v}{2}z^2 + \int_{|u|>0} (e^{izu} - 1 - i\phi(u)z)n(du),$$

$$m \in \boldsymbol{R}, \qquad v \geqq 0, \qquad \int_{|u|>0} (u^2 \wedge 1)n(du) < \infty.$$

証明 μ は無限可解であるから，定義により $\mu_1=\mu$ となる連続半群 $\{\mu_t\}$ がある．$\{\mu_t\}$ に対して上述の一様 Lévy 過程 X を考えると，上にのべたように $\mathcal{F}\mu_t(z)=\exp(t\psi(z))$ となる．ここで $t=1$ とおけば上の定理が得られる．∎

以上の考察から無限可解分布，連続半群 $\{\mu_t\}$，一様 Lévy 過程 X の間に対応関係がある．

$$\mu \underset{\mu=\mu_1}{\text{——}} \{\mu_t\} \underset{\mu_t=X(t)\ \text{の確率法則}}{\text{————————}} X$$

$X\to\{\mu_t\}$，$\{\mu_t\}\to\mu$ は明らかに全射である．$X\to\{\mu_t\}$ は定理 5.29 により法則同等な X を同一視すれば全単射である．$\{\mu_t\}\to\mu$ が全単射であることがいえれば，上の対応は完全な 1 対 1 対応となる．

$\{\mu_t\}\to\mu$ が全単射であることを証明しよう．$\{\mu_t\}\to\mu$，$\{\nu_t\}\to\mu$ とする．$\{\mu_t\}$，$\{\nu_t\}$ に対応して一様 Lévy 過程 X, Y を考え，上述の考察をすると

$$\mathcal{F}\mu_t(z) = e^{t\psi(z)}, \qquad \mathcal{F}\nu_t(z) = e^{t\theta(z)}$$

の形にかける．$\mu_1=\nu_1=\mu$ であるから，

$$e^{\psi(z)} = e^{\theta(z)},$$
$$(e^{\psi(z)/n})^n = (e^{\theta(z)/n})^n,$$
$$e^{\psi(z)/n} = e^{\theta(z)/n}\varepsilon_n(z), \qquad \varepsilon_n(z) = e^{i2\pi k/n} \quad (k=1, 2, \cdots, n).$$

$\psi(z), \theta(z)$ はともに前述の形にかける（m, v, n は異なるかもしれない）から z について連続である．積分の部分の連続性をいうには n に関する条件を考慮する必要があるが，平易な計算であるから，読者にまかせる．したがって

$$\varepsilon_n(z) = \frac{e^{\psi(z)/n}}{e^{\theta(z)/n}}$$

も z について連続である．しかし $\varepsilon_n(z)$ は離散値をとるから定数でなくてはならない．かくして $\varepsilon_n(z)=\varepsilon_n(0)=1$ が得られ，

$$e^{\psi(z)/n} = e^{\theta(z)/n}$$

がすべての n に対してなりたつ．したがって

$$\psi(z) = \lim_{n\to\infty} n(e^{\psi(z)/n}-1) = \lim_{n\to\infty} n(e^{\theta(z)/n}-1) = \theta(z)$$

が得られる．これから

$$\mathcal{F}\mu_t(z) = e^{t\psi(z)} = e^{t\theta(z)} = \mathcal{F}\nu_t(z) \quad \text{すなわち} \quad \mu_t = \nu_t$$

がすべての t に対してなりたつ．これは $\{\mu_t\}\to\mu$ が単射であることを意味する．μ が無限可解のときには，定義により $\mu=\mu_1$ となる連続半群 $\{\mu_t\}$ があるが，実はこの $\{\mu_t\}$ はただひとつしかないわけである．

無限分解可能な μ は $\mathcal{F}\mu(z)$ で特長づけられ，$\mathcal{F}\mu(z)$ に対しては上述の形の $\psi(z)$ が定まって $\mathcal{F}\mu(z)=\exp\{\psi(z)\}$ となり，さらに $\psi(z)$ は (n, m, v) によってあらわされる．しかしこの m, v, n は μ によって（法則同等を除いて）きまる一様 Lévy 過程 X によって（実は X の確率法則によって）きまるので，結局 (n, m, v) も μ によってきまることがわかる．かくして

$$(n, m, v) \text{ —— } \psi(z) \text{ —— } \mathcal{F}\mu(z) \text{ —— } \mu \text{ —— } \{\mu_t\} \text{ —— } X$$

の対応が完全に 1 対 1 となる．

ここで注意すべきことは，n と v は μ によってきまるが，m は $\phi(u)$ を固定したとき，きまるので，$\phi(u)$ をかえるとそれに応じて変動するということである．$\phi(u)$ としては，

$$\phi(u) = u + o(u^2), \qquad \phi(u) \text{ は有界}$$

なる任意のものをとってもよいことは前節例題 5.10 (i) で注意した．しかし測度 $n(du)$ が $u=0$ の近傍であまりに大きくないときには，$\phi(u)\equiv 0$ ととることもできる．例えば

$$\int_{0<|u|<1} |u| n(du) < \infty$$

のときには $\phi(u)\equiv 0$ としてよくて，

$$\psi(z) = imz - \frac{v}{2}z^2 + \int_{|u|>0} (e^{izu}-1)n(du)$$

となる．この m が前の m と異なることはいうまでもない．この特別の場合として，n が $\boldsymbol{R}-\{0\}$ 全体で有界である場合がある．さらに $m=0$，$v=0$ のときには

$$\psi(z) = \int_{|u|>0} (e^{izu}-1)n(du)$$

となる．これに対応する μ を**複合 Poisson 分布**という．$n(\boldsymbol{R}-\{0\})$ を λ とする．$\lambda=0$ のときは $\mu=\delta$ となり，つまらない．このばあいを除いて，複合 Poisson 分布の構造をしらべて見よう．$\nu(E)=n(E)/\lambda$ は確率分布であって，

$$\psi(z)=\lambda\int(e^{izu}-1)\nu(du)=\lambda(\mathscr{F}\nu(z)-1),$$

$$\mathscr{F}\mu(z)=\exp\{-\lambda+\lambda\mathscr{F}\nu(z)\}=\sum_{n=0}^{\infty}e^{-\lambda}\frac{\lambda^n(\mathscr{F}\nu(z))^n}{n!}$$

$$=\sum_{k=1}^{\infty}e^{-\lambda}\frac{\lambda^n}{n!}(\mathscr{F}\nu^{*n})(z),$$

$$\nu^{*n}=\nu*\nu*\cdots*\nu \quad (n\text{ 個}),\quad \nu^{*0}=\delta$$

したがって

$$\mu=\sum_{n=1}^{\infty}e^{-\lambda}\frac{\lambda^n}{n!}\nu^{*n}.$$

すなわち μ は ν^{*n} を Poisson 分布 p_λ の重みで平均したものである．この複合 Poisson 分布を $p_{\lambda,\nu}$ であらわし，対応する Lévy 過程を**複合 Poisson 過程**という．

例題 5.11 （i） 無限可解な μ の台が半直線 $[0,\infty)$ にのっているときには，$n(du)$ も $[0,\infty)$ にのっていて

$$\int_{0+}^{\infty}(u\wedge 1)n(du)<\infty$$

となり，

$$\psi(z)=imz+\int_{0+}^{\infty}(e^{izu}-1)n(du) \qquad (m\geqq 0)$$

とかける．μ の台が $[0,\infty)$ 上にあることと対応する一様 Lévy 過程 X が増大（見本関数がほとんど確実に増大）であることとは同等である．

［ヒント］ $\{\mu_t\}$ を μ に対応する半群とすれば，

$$\mu_t*\mu_{1-t}=\mu_1=\mu, \qquad 0\leqq t\leqq 1$$

により，すべての $t<1$ に対して μ_t の台は $[0,\infty)$ に入る．これからすべての t に対してこのことがいえる．さて $t>s$ のとき

$$P\{X(t)-X(s)\geqq 0\}=\mu_{t-s}([0,\infty))=1$$

であるから，

$$P\Big(\bigcap_{r<r'}\{X(r')\geqq X(r)\}\Big)=1 \qquad (r,r'\in\boldsymbol{Q}).$$

X は D 過程であるから

$$P\Big(\bigcap_{s<t}\{X(s)\leqq X(t)\}\Big)=1$$

となり，$X(t,\omega)$ が t の増加関数となる確率は1に等しい．逆は明らかであるから，定理の後半は証明された．

$X(t,\omega)$ は t の増加関数であるから，これを連続部分と飛躍の和の部分にわけると

$$X(t,\omega)=G(t,\omega)+\int_0^t\int_{0+}^{\infty}uN(dsdu)$$

となる．N は X の飛躍の Poisson 配置である．G は連続増大の Lévy 過程であるから，Gauss 型である．したがって $G(t,\omega)\geqq 0$ a.s. でなければならないが，$G(t,\omega)$ は Gauss 分布に従うから，$G(t,\omega)=m(t)\geqq 0$（ω に無関係）でなければならない．$\{X(t+s)-X(s), t\in[0,\infty)\}$ が X と法則同等であることから $m(t)=m\cdot t$ $(m\geqq 0)$ が得られ，同時に

$$E(N(dudt))=dtn(du)$$

となる．これから定理の前半が得られるが，その際

$$\int_{0+}^{1}un(du)<\infty$$

がなりたつことを注意する必要がある．このためには補題 5.4 を用い

$$E\Big(\exp\Big\{-\alpha\int_{s\leqq 1}\int_{1/n\leqq u<1}uN(dsdu)\Big\}\Big)\geqq E(e^{-\alpha X(1)})$$

を示し，左辺が

$$\exp\Big\{\int_0^1 ds\int_{1/n}^{1}(e^{-\alpha u}-1)n(du)\Big\}$$

に等しいことに注意して，まず $n\to\infty$，つぎに $\alpha\downarrow 0$ として見ればよい．（$\beta>0$ が十分小さいときには $e^{-\beta}-1<-\beta/2$ となることに注意せよ．）

（ii） N が Poisson 分布 p_λ に従い，$X_1, X_2, \cdots$ が同分布 ν に従い，しかも $N, X_1, X_2, \cdots$ が独立とする．Y を

$$Y=X_1+X_2+\cdots+X_N$$

と定義すると，Y は複合 Poisson 分布 $p_{\lambda,\nu}$ に従う．

［ヒント］ $$E(e^{izY}) = \sum_{n=0}^{\infty} E(\exp(iz(X_1+X_2+\cdots+X_n)), N=n)$$
$$= \sum_{n=0}^{\infty} P\{N=n\} \prod_{k=1}^{n} E(\exp(izX_k))$$
$$= \sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}(\mathcal{F}\nu(z))^n = \exp\{-\lambda+\lambda\mathcal{F}\nu(z)\}$$
$$= \exp\left(\lambda\int(e^{izu}-1)\nu(du)\right).$$

(iii) 瞬間毎に独立に dt の間に λdt の確率で事故がおこり，事故のおこったときの損害額は分布 ν に従うとする．事故の生起も損害額もすべて独立であるとするとき，時点 t までの損害総額 $X(t,\omega)$ は‘複合 Poisson 分布 $p_{\lambda,\nu}$ に対応する一様 Lévy 過程（これを複合 Poisson 過程という）’となることを説明せよ．

(iv) Poisson 分布 p_λ，Cauchy 分布 $C_{m,c}$ の Lévy 分解を定めよ．（これに対する一様 Lévy 過程を **Poisson 過程**，**Cauchy 過程**という．）

［ヒント］ p_λ に対しては $\psi(z)=\int_{0+}^{\infty}(e^{izu}-1)n(du)$，$n=\lambda\cdot\delta_1$，$C_{m,c}$ に対しては
$$\mathcal{F}C_{m,c} = \exp(imz-c|z|)$$
に注意し，つぎの式を利用して $|z|$ をかきなおせ．
$$|z| = \frac{2}{\pi}\int_0^{\infty}\frac{1-\cos zu}{u^2}du$$
$$= \lim_{n\to\infty}\frac{1}{\pi}\int_{|u|>1/n}\frac{1-e^{izu}-izu}{u^2}du.$$

§5.12 Markov 過程と転移確率

$\{X_0, X_1, X_2, \cdots\}$ を離散時変数の確率過程とする．説明をわかりやすくするため，X_n のとる値も整数のばあいから始める．$X_0, X_1, \cdots, X_n$ までの値が $a_0, a_1, \cdots, a_n$ であったときの X_{n+1} の（条件付）確率法則
$$P\{X_{n+1}=a_{n+1}\,|\,X_0=a_0, X_1=a_1, \cdots, X_n=a_n\}$$
と，X_n の値が a_n であったときの X_{n+1} の確率法則
$$P\{X_{n+1}=a_{n+1}\,|\,X_n=a_n\}$$
とは一般にことなる．この両者が $a_0, a_1, \cdots, a_{n+1}$ のとり方のいかんにかかわらず

一致するとき，X は **Markov 過程**であるという．または X は **Markov 性**をもつという．

たとえば $X_1, X_2, \cdots$ が独立なときには，両者はともに $P\{X_{n+1}=a_{n+1}\}$ に等しいから，$X=\{X_n\}$ は Markov 性をもつ．Markov 性は独立性よりも一般な性質である．$\{X_n\}$ が独立のときに，
$$Y_n = \sum_{k=0}^{n} X_k, \quad n=0,1,2,\cdots$$
は明らかに独立ではないが，Markov 性をもつ．実際
$$P\{Y_{n+1}=a_{n+1}\,|\,Y_0=a_0, Y_1=a_1, \cdots, Y_n=a_n\} = P\{X_{n+1}=a_{n+1}-a_n\},$$
$$P\{Y_{n+1}=a_{n+1}\,|\,Y_n=a_n\} = P\{X_{n+1}=a_{n+1}-a_n\}$$
が条件付確率の定義と $\{X_n\}$ の独立性から証明され，両者は一致する．

壺の中に赤玉 r，黒玉 b，合計 $N=r+b$ 個の玉が入っているとする．これから玉を1個ずつとりだし，n 回目にでた玉が赤か黒かに応じて $X_n=1,0$ とする．このとき，$X_1, X_2, \cdots, X_N$ は独立性も Markov 性も持たない．実際
$$P\{X_{n+1}=1\,|\,X_1=a_1, X_2=a_2, \cdots, X_n=a_n\} = \frac{r-a_1-a_2-\cdots-a_n}{N-n},$$
$$P\{X_{n+1}=1\,|\,X_n=a_n\} = \frac{r-a_n}{N-1}$$
で両者は異なる．しかし
$$Y_n = X_1+X_2+\cdots+X_n = n \text{ 回までにでた赤玉の総数}$$
とおくと $a_{n+1}=a_n$，a_n+1 に応じて
$$P\{Y_{n+1}=a_{n+1}\,|\,Y_1=a_1, Y_2=a_2, \cdots, Y_n=a_n\}$$
$$= P\{Y_{n+1}=a_{n+1}\,|\,Y_n=a_n\} = \frac{b-(n-a_n)}{N-n}, \quad \frac{r-a_n}{N-n}$$
となり，$\{Y_n\}$ は Markov 性をもつ．

さて X_n が一般の実確率変数のときには，上の Markov 性の定義はそのままでは意味がなくなる．上の定義を念頭において，任意の $E\in\mathcal{B}^1$ に対し
$$P\{X_{n+1}\in E\,|\,(X_1, X_2, \cdots, X_n)\} = P\{X_{n+1}\in E\,|\,X_n\} \text{ a.s.}$$
すなわち
$$P\{X_{n+1}\in E\,|\,\sigma[X_1, X_2, \cdots, X_n]\} = P\{X_{n+1}\in E\,|\,\sigma[X_n]\} \text{ a.s.}$$
のとき，$\{X_n\}$ は **Markov 性**をもつ．または $\{X_n\}$ は **Markov 過程**であるという．

この定義から

$$P\{X_{n+m}\in E\mid(X_1,X_2,\cdots,X_n)\}=P\{X_{n+m}\in E\mid X_n\}\quad\text{a.s.}$$

がでる．これをいうために，まず Markov 性の定義から，任意の有界 Borel 関数 $f: \boldsymbol{R}\to\boldsymbol{R}$ に対し

$$E\{f(X_{n+1})\mid(X_1,X_2,\cdots,X_n)\}=E\{f(X_{n+1})\mid X_n\}\quad\text{a.s.}$$

がでることに注意しておく．あるいは $\sigma[X_{n+1}]$ に関して可測な有界確率変数 Y_{n+1} に対して

$$E\{Y_{n+1}\mid(X_1,X_2,\cdots,X_n)\}=E\{Y_{n+1}\mid X_n\}\quad\text{a.s.}$$

といってもよい．さて証明すべき式は $m=1$ のときには定義そのものである．m のときになりたつならば(以下 a.s. を省略することがある)

$$\begin{aligned}&P\{X_{n+m+1}\in E\mid(X_1,X_2,\cdots,X_n)\}\\&=E\{P\{X_{n+m+1}\in E\mid(X_1,X_2,\cdots,X_{n+m})\}\mid(X_1,X_2,\cdots,X_n)\}\\&\qquad(\because\ \sigma[X_1,X_2,\cdots,X_{n+m}]\supset\sigma[X_1,X_2,\cdots,X_n])\\&=E\{P\{X_{n+m+1}\in E\mid X_{n+m}\}\mid(X_1,X_2,\cdots,X_n)\}.\end{aligned}$$

$P\{X_{n+m+1}\in E\mid X_{n+m}\}$ は $\sigma[X_{n+m}]$ に関し可測かつ有界であるから，m に関する仮定により最後の式は

$$E\{P\{X_{n+m+1}\in E\mid X_{n+m}\}\mid X_n\}$$

に等しい．したがって

$$P\{X_{n+m+1}\in E\mid(X_1,X_2,\cdots,X_n)\}=E\{P\{X_{n+m+1}\in E\mid X_{n+m}\}\mid X_n\}.$$

右辺は $\sigma[X_n]$ に関して可測であるから，左辺も同様で，

$$\begin{aligned}P\{X_{n+m+1}\in E\mid(X_1,X_2,\cdots,X_n)\}&=E\{P\{X_{n+m+1}\in E\mid(X_1,X_2,\cdots,X_n)\}\mid X_n\}\\&=P\{X_{n+m+1}\in E\mid X_n\}\end{aligned}$$

となる．したがって Markov 性から

$$P\{X_{n+m}\in E\mid(X_1,X_2,\cdots,X_n)\}=P\{X_{n+m}\in E\mid X_n\}\quad\text{a.s.}$$

が示された．この式で $m=1$ とおけば，Markov 性の定義が得られるから，Markov 性を上の式で定義してもよい．この定義は連続時変数のときに Markov 性の定義の方法を暗示している．連続時変数のときにはつぎの時点というのが意味がないから，もとの定義はそのまま用いることができない．

$\{X_t\}$ を連続時変数の確率過程とする．

$$P\{X_u\in E\mid X_s,s\leqq t\}=P\{X_u\in E\mid X_t\}\quad\text{a.s.},\qquad u>t$$

のとき，$\{X_t\}$ は **Markov 性**をもつ(**Markov 過程**である)という．これはまた任意の有界 Borel 関数 $f: \boldsymbol{R}\to\boldsymbol{R}$ に対して

$$E\{f(X_u)\mid X_s,s\leqq t\}=E\{f(X_u)\mid X_t\}\quad\text{a.s.},\qquad u>t$$

と定義してもよい．実はすべての有界 Borel 関数に対して仮定しなくても，$f\in C_0^\infty=C_0^\infty(\boldsymbol{R})$ (＝無限回可微分でコンパクトな台をもつ関数の族)に対して上の式がなりたつことを要求しておけば，十分であることは，確率測度の定義から明らかである．

$P\{X(u)\in E\mid X(t)=x\}$ を $p(t,x,u,E)$ であらわし，Markov 過程 $\{X_t\}$ の(時点 t から u への)**転移確率**(transition probability)という．この $p(t,x,u,E)$ は t,u,x をきめたとき，$E\in\mathcal{B}^1$ については確率法則となり，t,u,E をきめたとき，x について Borel 関数となるように定めることができる(§3.6)．このきめ方は一意ではない．実際二通り $p(t,x,u,E)$，$q(t,x,u,E)$ あるとすれば

$$p(t,x,u,E)=q(t,x,u,E)$$

がほとんどすべて$(P^{X(t)})$の x に対してなりたつ．この例外集合は E には無関係にとれるが，t,u に無関係にとることは一般にできない．$s<t<u$ のとき

$$\begin{aligned}p(s,X(s),u,E)&=P\{X(u)\in E\mid X(s)\}\\&=E\{P\{X(u)\in E\mid X(\theta),\theta\leqq t\}\mid X(s)\}\qquad(t>s\text{ に注意})\\&=E\{P\{X(u)\in E\mid X(t)\}\mid X(s)\}\\&=E\{p(t,X(t),u,E)\mid X(s)\}\\&=\int p(t,y,u,E)P\{X(t)\in dy\mid X(s)\}\\&=\int p(t,y,u,E)p(s,X(s),t,dy).\end{aligned}$$

したがって **Chapman-Kolmogorov 方程式**：

$$\text{(CK)}\qquad p(s,x,u,E)=\int_{\boldsymbol{R}}p(s,x,t,dy)p(t,y,u,E)$$

がほとんどすべての$(P^{X(s)})$の x に対してなりたつ．例外の $P^{X(s)}$ 零集合は E には無関係にとれるが，s,t,u には関係する．

Chapman-Kolmogorov 方程式に例外点があることは上述した通りであるが，転移確率 $p(s,x,t,E)$ を適当にとって，この方程式が例外点なしになりたつよう

にできるとき，このようなよい性質の $p(s,x,t,E)$ だけを Markov 過程 $\{X(t)\}$ の転移確率とよぶことにする．この意味では転移確率は必ずしも存在しないが，多くの有用な Markov 過程では存在する．

Markov 過程から離れて，一般に $s,t\in[0,\infty)$ $(s<t)$，$x\in\boldsymbol{R}$，$E\in\mathcal{B}^1$ の関数 $p(s,x,t,E)$ がつぎの3条件をみたすときに**転移確率**という．

(T.1) s,t,E を固定したとき，x に関して Borel 可測，

(T.2) s,t,x を固定したとき，E に関して確率測度，

(T.3) すべての $0\leqq s<t<u$，$x\in\boldsymbol{R}$，$E\in\mathcal{B}^1$ に対し Chapman-Kolmogorov の方程式

$$p(s,x,u,E)=\int_R p(s,x,t,dy)p(t,y,u,E)$$

がなりたつ．

転移確率 $p(s,x,t,E)$ が Markov 過程 $\{X(t)\}$ の転移確率であるとは，すべての s,t,E に対し

$$p(s,X(s),t,E)=P\{X(t)\in E\mid X(s)\}\ \text{a.s.}(P)$$

がなりたつことである．これはまたすべての s,t,E に対し

$$p(s,x,t,E)=P\{X(t)\in E\mid X(s)=x\}\ \text{a.s.}(P^{X(s)})$$

といってもよい．$\{X(t)\}$ が Markov 過程であることがわからないときにも

$$P\{X(t)\in E\mid X(\theta),\theta\leqq s\}=p(s,X(s),t,E)\ \text{a.s.}(P)$$

のときには，$\{X(t)\}$ は Markov 過程となり，$p(s,x,t,E)$ はその転移確率となる．実際，上の式がなりたてば，右辺が $\sigma[X(s)]$ について可測であるから

$$\text{左辺}=E\{P\{X(t)\in E\mid X(\theta),\theta\leqq s\}\mid X(s)\}=P\{X(t)\in E\mid X(s)\}$$

となり，$\{X(t)\}$ は Markov，しかも

$$P\{X(t)\in E\mid X(s)\}=p(s,X(s),t,E)\ \text{a.s.}(P)$$

で $p(s,x,t,E)$ は $\{X(t)\}$ の転移確率である．

$p(s,x,t,E)$ を転移確率とする．もし

$$p(s,x,t,E)=p(0,x,t-s,E)$$

がつねになりたつときには $p(s,x,t,E)$ を（時間的に）**一様な転移確率**といい，$p(t-s,x,E)$ であらわす．このときには Chapman-Kolmogorov 方程式は

$$p(t+s,x,E)=\int_R p(t,x,dy)p(s,y,E)$$

となる．一様な転移確率をもつ Markov 過程を**一様 Markov 過程**という．

Markov 性は

$$P\{X(t)\in E\mid X(\theta),\theta\leqq s\}=P\{X(t)\in E\mid X(s)\},\qquad t>s$$

によって定義されるが，これを検証するために，つぎの定理が有用である．

定理 5.31 $X=\{X(t)\}$ が Markov 性をもつためには，任意の $0\leqq s_1<s_2<\cdots<s_n<t$ に対し

$$P\{X(t)\in E\mid X(s_1),X(s_2),\cdots,X(s_n)\}=P\{X(t)\in E\mid X(s_n)\}$$

がなりたつことが必要十分である．

証明 $s<t$ を任意に固定し，$[0,s]$ の分点系

$$\lambda=\{0\leqq s_1<s_2<\cdots<s_n=s\}$$

の全体を Λ であらわし，上の λ に対し

$$\mathcal{B}_\lambda=\sigma[X(s_1),X(s_2),\cdots,X(s_n)]$$

とおく．$\lambda\subset\lambda'$ のとき $\mathcal{B}_\lambda\subset\mathcal{B}_{\lambda'}$ であるから，$\mathcal{B}_\lambda$，$\lambda\in\Lambda$ は有向集合である．仮定により

$$P\{X(t)\in E\mid\mathcal{B}_\lambda\}=P\{X(t)\in E\mid X(s)\},\qquad \lambda\in\Lambda.$$

下に示す補題により

$$P\left\{X(t)\in E\,\middle|\,\bigvee_\lambda\mathcal{B}_\lambda\right\}=P\{X(t)\in E\mid X(s)\}.$$

$\sigma[X(\theta),\theta\leqq s]=\bigvee_\lambda\mathcal{B}_\lambda$ であるから，

$$P\{X(t)\in E\mid X(\theta),\theta\leqq s\}=P\{X(t)\in E\mid X(s)\}.\qquad\blacksquare$$

補題 5.10 $\mathcal{B}_\lambda$，$\lambda\in\Lambda$ が有向系とする．

$$P\{A\mid\mathcal{B}_\lambda\}=Y,\ \ \lambda\in\Lambda\ \ (Y\text{ は }\lambda\text{ に無関係})\Longrightarrow P\left\{A\,\middle|\,\bigvee_\lambda\mathcal{B}_\lambda\right\}=Y.$$

証明 仮定により

$$P(A\cap C)=E(Y,C),\qquad C\in\mathcal{B}_\lambda,\ \ \lambda\in\Lambda.$$

したがって任意の $C\in\bigcup_\lambda\mathcal{B}_\lambda$（集合和）に対し，上の等式がなりたつ．$\mathcal{B}_\lambda$，$\lambda\in\Lambda$ の有向性により，$\bigcup_\lambda\mathcal{B}_\lambda$ は乗法族である．ゆえに Dynkin 族定理を用いて，上の式が $\sigma\left[\bigcup_\lambda\mathcal{B}_\lambda\right]$ すなわち $\bigvee_\lambda\mathcal{B}_\lambda$ の中の任意の集合 C に対してもなりたつことがわかる．これは $P\left\{A\,\middle|\,\bigvee_\lambda\mathcal{B}_\lambda\right\}=Y$ を意味する．（$P\{A\mid\mathcal{B}_\lambda\}=Y$ により，Y は可測

$\mathcal{B}_\lambda$，したがって当然可測 $\bigvee_\lambda \mathcal{B}_\lambda$ であることに注意せよ.) ∎

$p(s, x, t, E)$ を転移確率とする. $X=\{X(t)\}$ が $p(s, x, t, E)$ を転移確率とする Markov 過程であるためには

$$P\{X(t)\in E\,|\,X(s_1), X(s_2), \cdots, X(s_n)\} = p(s_n, X(s_n), t, E),$$
$$0\leqq s_1<s_2<\cdots<s_n<t$$

を検証すればよいことを注意しよう. 実際この式がなりたてば

$$\begin{aligned}P\{X(t)\in E\,|\,X(s_n)\} &= E\{P\{X(t)\in E\,|\,X(s_1), X(s_2), \cdots, X(s_n)\}\,|\,X(s_n)\}\\ &= E\{p(s_n, X(s_n), t, E)\,|\,X(s_n)\}\\ &= p(s_n, X(s_n), t, E).\end{aligned}$$

したがって，上の仮定の式と合わせて

$$\begin{aligned}P\{X(t)\in E\,|\,X(s_1), X(s_2), \cdots, X(s_n)\} &= P\{X(t)\in E\,|\,X(s_n)\}\\ &= p(s_n, X(s_n), t, E),\\ &0\leqq s_1<s_2<\cdots<s_n<t.\end{aligned}$$

上の定理により X は Markov 過程で，しかも $p(s, x, t, E)$ はその転移確率である.

定理 5.32 $X=\{X(t)\}$ を加法過程とし，

$$\mu_{s,t}(E) = P\{X(t)-X(s)\in E\},\qquad s<t$$

とおくと，X は

$$p(s, x, t, E) = \mu_{s,t}(E-x)$$

を転移確率とする Markov 過程である.

証明 $p(s, x, t, E)$ が転移確率の性質をもっていることは

$$\mu_{s,t}*\mu_{t,u} = \mu_{s,u}\qquad (s<t<u)$$

から明らかである. したがって，任意の

$$0\leqq s_1<s_2<\cdots<s_n<t$$

に対し，

$$P\{X(t)\in E\,|\,X(s_1), X(s_2), \cdots, X(s_n)\} = \mu_{s_n,t}(E-X(s_n))$$

をいえば，上の注意により，この定理が正しいことがわかる.

X の加法性により

$$\boldsymbol{X} = (X(s_1), X(s_2), \cdots, X(s_n)),\qquad Y = X(t)-X(s_n)$$

は独立であるから，例題 3.6 (ii) の第2式 (で $s=\boldsymbol{X}$ とおいた式) により

$$E(f(\boldsymbol{X}, Y)\,|\,\boldsymbol{X}) = E(f(\boldsymbol{x}, Y))\,|_{\boldsymbol{x}=\boldsymbol{X}}\qquad (\boldsymbol{x}=(x_1, x_2, \cdots, x_n)).$$

ここで

$$f(\boldsymbol{x}, y) = 1_E(\pi_n(\boldsymbol{x})+y)\qquad (\pi_n : \boldsymbol{x}=(x_1, x_2, \cdots, x_n)\to x_n)$$

とおくと

$$\begin{aligned}&P\{\pi_n(\boldsymbol{X})+Y\in E\,|\,\boldsymbol{X}\}\\ &= P\{\pi_n(\boldsymbol{x})+Y\in E\}\,|_{\boldsymbol{x}=\boldsymbol{X}} = P\{Y\in E-\pi_n(\boldsymbol{x})\}\,|_{\boldsymbol{x}=\boldsymbol{X}}\\ &= \mu_{s_n,t}(E-\pi_n(\boldsymbol{x}))\,|_{\boldsymbol{x}=\boldsymbol{X}} = \mu_{s_n,t}(E-X(s_n)).\end{aligned}$$

$\pi_n(\boldsymbol{X})+Y=X(s_n)+X(t)-X(s_n)=X(t)$ であるから，上の式は目的の等式に外ならない. ∎

転移確率 $p(s, x, t, E)$ をもつ Markov 過程 $X=\{X(t)\}$ があるとき，任意の

$$0\leqq s_1<s_2<\cdots<s_n,\qquad f\colon \boldsymbol{R}^n\longrightarrow \boldsymbol{R}\ \ (\text{有界 Borel 可測})$$

に対して

$$\begin{aligned}&E\{f(X(s_1), X(s_2), \cdots, X(s_n))\}\\ &= \int_R\int_R\cdots\int_R p_{s_1}(dx_1)p(s_1, x_1, s_2, dx_2)\cdots p(s_{n-1}, x_{n-1}, s_n, dx_n)f(x_1, x_2, \cdots, x_n)\end{aligned}$$

がなりたつ. ここに $p_{s_1}(E)=P\{X(s_1)\in E\}$ とする. これを証明するには，$f(x_1, x_2, \cdots, x_n)=f_1(x_1)f_2(x_2)\cdots f_n(x_n)$ のときに証明すれば十分である. この特別の場合には

$$\begin{aligned}\text{左辺} &= E[f_1(X(s_1))f_2(X(s_2))\cdots f_n(X(s_n))]\\ &= E[f_1(X(s_1))f_2(X(s_2))\cdots f_{n-1}(X(s_{n-1}))E(f_n(X(s_n))\,|\,X(s_1), X(s_2),\\ &\qquad \cdots, X(s_{n-1}))]\\ &= E[f_1(X(s_1))f_2(X(s_2))\cdots f_{n-1}(X(s_{n-1}))E(f_n(X(s_n))\,|\,X(s_{n-1}))]\\ &\qquad (\text{Markov 性})\\ &= E[f_1(X(s_1))f_2(X(s_2))\cdots f_{n-1}(X(s_{n-1}))\\ &\qquad \cdot\int p(s_{n-1}, X(s_{n-1}), s_n, dy)f_n(y)].\end{aligned}$$

最後の2因子の積は $X(s_{n-1})$ の関数であるから，再び上の変形を行い，これをくりかえして

$$\text{左辺} = \int_R\int_R\cdots\int_R p_{s_1}(dx_1)p(s_1, x_1, s_2, dx_2)$$

$$\cdots p(s_{n-1}, x_{n-1}, s_n, dx_n)f_1(x_1)f_2(x_2)\cdots f_n(x_n).$$

上の事実は Markov 過程 X はその転移確率 $p(s, x, t, E)$ と $p_0(E)=P\{X_0 \in E\}$ (これを X の**初期分布**という)によって法則同等を除いて一意に定まることがわかる．逆に分布 p_0 と転移確率 $p(s, x, t, E)$ とが与えられたとき，これに対応する Markov 過程の存在を無条件でいうのには，基礎の確率空間 (Ω, P) が完全可分であるという本講の制限は強すぎる．この制限を除いて，かつ Kolmogorov の拡張定理を $\boldsymbol{R}^\infty$ の上から $\boldsymbol{R}^A$ (A は任意の集合，ここのばあいには $A=[0, \infty)$) まで拡張しておく(これは可能)必要がある．しかし応用上は，このような一般のばあいを考える必要はおこらない．

例題 5.12 (i) $X=\{X(t)\}$ のとる値が可算個，例えば $\{1, 2, 3, \cdots\}$ とすれば，その転移確率 $p(s, i, t, E)$ は E が 1 点集合 $\{j\}$ のときだけを与えておけば，十分である．$p(s, i, t, \{j\})$ を $p_{s,t}(i, j)$ であらわし，行列

$$p_{s,t} = \{p_{s,t}(i, j) \mid i, j=1, 2, \cdots\}$$

を**転移行列**という．このように考えると，Chapman-Kolmogorov の方程式は行列の掛算で

$$p_{s,t}\cdot p_{t,u} = p_{s,u}$$

とかけることを示せ．

(ii) $X=\{X(t), t\geqq 0\}$ を $p(s, x, t, E)$ を転移確率とする Markov 過程とするとき，$X^{(a)}=\{X(t), t\geqq a\}$ を $(\Omega, P(\cdot \mid X(a)))$ の上で考えて，ほとんど確実に $p(s, x, t, E)$ $(t>s\geqq a)$ を転移確率とする Markov 過程であることを証明せよ．

[ヒント] Markov 性の定義と，条件付確率法則 $P(\cdot \mid X(s))$ の意味をよく考えよ．

(iii) $X=\{X(t)\}$ は Markov 過程であるとき，$Y=f(X(t))$ (f は単射) は Markov 過程となることを示せ．f が単射でないときには，この事実は必ずしもなりたたないことを反例をあげて説明せよ．

(iv) $B=\{B(t)\}$ が Wiener 過程(したがってもちろん Markov 過程)のとき，$Y=\{B(t)^2\}$ は Markov 過程となることを示せ．この事実と前問の反例との関係を考察せよ．

[ヒント] B の転移確率 $p(s, x, t, E)=N(t-s, E-x)$ が

$$p(s, -x, t, -E) = p(s, x, t, E), \qquad -E = \{-y \mid y \in E\}$$

をみたす点に注意せよ．

§5.13 生成作用素

$B(\boldsymbol{R})$ で $\boldsymbol{R}$ の上の有界 Borel 関数の全体をあらわす．任意の転移確率 $p(s, x, t, E)$ に対し，

$$(p_{s,t}f)(x) = \int_{\boldsymbol{R}} p(s, x, t, dy)f(y)$$

とおくと，$f\in B(\boldsymbol{R})$ のときには $p_{s,t}f$ が定義され，しかも $p(s, x, t, E)$ が x に関して Borel 可測であることから，$p_{s,t}f \in B(\boldsymbol{R})$．したがって

$$p_{s,t}\colon B(\boldsymbol{R}) \longrightarrow B(\boldsymbol{R}).$$

$p_{s,t}$ を**転移作用素**という．逆に $p_{s,t}$ からは $p(s, x, t, E)$ が $p_{s,t}1_E$ として得られるから，転移確率のかわりに転移作用素 $p_{s,t}$ を考えることができる．転移確率に関する Chapman-Kolmogorov の方程式は $p_{s,t}$ における(作用素としての)結合関係

$$p_{s,t}p_{t,u} = p_{s,u}$$

であらわされる．

すべての s, x に対し

$$p(s, x, s, E) = \delta_x(E), \qquad \lim_{h\downarrow 0} p(s, x, s+h, \cdot) = \delta_x(\cdot) \quad (\text{弱収束})$$

のとき，この転移確率は**右連続**であるという．これは対応する転移作用素の言葉でいえば

$$p_{s,s} = I \qquad (\text{恒等作用素}),$$

$$\lim_{h\downarrow 0} p_{s,s+h}f(x) = f(x) \qquad (f\ \text{は任意の有界連続関数})$$

といってもよい．このとき $p_{s,t}$ は**右連続**という．

もし $p(s, x, t, E)$ がある Markov 過程 $X=\{X_t\}$ の転移確率であれば，$p(s, x, t, E)$ の右連続性から，X の確率右連続性がでる．実際，有界正値連続関数 $f(x)$ を $f(0)=0$，$f(x)=1$ $(|x|>\varepsilon)$ となるようにとると，

$$P\{|X(s+h)-X(s)|>\varepsilon\} \leqq E\{f(X(s+h)-X(s))\}$$

$$= \int_{\boldsymbol{R}} P\{X(s) \in dx\} \int_{\boldsymbol{R}} p(s, x, s+h, dy)f(y-x).$$

第2の積分は，$h\downarrow 0$ のとき $f(0)$ に収束し，しかも絶対値において $\|f\|_\infty=\sup_x |f(x)|$ でおさえられているから，有界収束定理により，上の2重積分は0に収束する．これは $\{X(t)\}$ が確率右連続であることを意味する．

関数の微分係数を考えるように，

$$A_s=\lim_{h\downarrow 0}\frac{p_{s,s+h}-I}{h}$$

を考えよう．ここで極限の定義をはっきりさせる必要がある．任意の

$f\in C_0^\infty(\boldsymbol{R})$ （＝コンパクトな台をもつ無限回可微分関数の族）

に対し

$$\lim_{h\downarrow 0}\frac{1}{h}(p_{s,s+h}f(x)-f(x)),\qquad x\in\boldsymbol{R}$$

が存在するとき，この極限を $(A_sf)(x)$ とあらわす．A_sf は Borel 可測関数ではあるが，必ずしも有界ではない．しかし A_s は線型作用素であることは明らかである．このように定義した A_s を転移作用素 $p_{s,t}$ の（s における）生成作用素という．また A_s を転移確率 $p(s,x,t,E)$ の生成作用素ということもあり，$p(s,x,t,E)$ が Markov 過程 $X=\{X(t)\}$ の転移確率であれば，X の生成作用素ともいう．

とくに $p(s,x,t,E)$ が一様転移確率であれば，これを $p(t-s,x,E)$ の形にかけることは前節でのべた．このときには対応する転移作用素 $p_{s,t}$ も p_{t-s} とかけて，

$$p_tf(x)=\int_R p(t,x,dy)f(y)$$

となる．したがって生成作用素 A_s があれば，これは s には無関係で A とかくことができて

$$Af(x)=\lim_{h\downarrow 0}\frac{1}{h}(p_hf(x)-f(x))$$

となる．したがって一様な Markov 過程の生成作用素 A_s も s に無関係である．

$B=\{B(t)\}$ を Wiener 過程とすると，B は加法過程であるから，もちろん Markov 過程で，

$$p(s,x,t,E)=N_{0,t-s}(E-x)$$

を転移確率とする．したがってもちろん一様である．その生成作用素 A を求めよう．$N(\boldsymbol{R})=1$ に注意して，

$$\begin{aligned}(Af)(x)&=\lim_{h\downarrow 0}\frac{1}{h}\left[\int_R N_{0,h}(dy-x)f(y)-f(x)\right]\\&=\lim_{h\downarrow 0}\frac{1}{h}\int_R N_{0,h}(dy)(f(x+y)-f(x))\\&=\lim_{h\downarrow 0}\frac{1}{h}\int_R N(dy)[f(x+\sqrt{h}\,y)-f(x)]\qquad N=N_{0,1}.\end{aligned}$$

$f\in C_0^\infty(\boldsymbol{R})$ とすれば，

$$f(x+\sqrt{h}\,y)-f(x)=\sqrt{h}\,yf'(x)+\frac{h}{2}y^2f''(x)+h^{3/2}y^3\theta.$$

ここで θ は x,y,h に関係するが，有界である．

$$\int_R yN(dy)=0,\qquad \int_R y^2N(dy)=1,\qquad \int_R y^3N(dy)=3$$

により，

$$\int_R N(dy)(f(x+\sqrt{h}\,y)-f(x))=\frac{h}{2}f''(x)+3h^{3/2}\theta_1(x,y,h)\qquad (\theta_1\text{: 有界}).$$

これから Af は $f\in C_0^\infty(\boldsymbol{R})$ に対し存在し

$$(Af)(x)=\frac{1}{2}f''(x)\quad \text{すなわち}\quad A=\frac{1}{2}\frac{d^2}{dx^2}.$$

したがって Wiener 過程の生成作用素は2階の微分作用素である．この微分作用素の係数は x に関係しない．

つぎに

$$X_t=B_t^3$$

を考えて見よう．$f(x)=x^3$ は単射であるから，例題 5.12 (iii) により，$\{X_t\}$ も Markov 過程であって，その転移作用素は

$$\begin{aligned}p_{s,t}f(x)&=\int_R f(y)P\{X(t)\in dy\,|\,X(s)=x\}=E\{f(X(t))\,|\,X(s)=x\}\\&=E\{f(B(t)^3)\,|\,B(s)=x^{1/3}\}=\int_R f(y^3)N_{0,t-s}(dy-x^{1/3})\\&=\int_R f((y+x^{1/3})^3)N_{0,t-s}(dy)=\int_R f((\sqrt{t-s}\,y+x^{1/3})^3)N(dy),\end{aligned}$$

$$p_{s,s+h}f(x)-f(x)=\int_R\{f((\sqrt{h}\,y+x^{1/3})^3)-f(x)\}N(dy),$$

$$f((\sqrt{h}\,y+x^{1/3})^3) = f(x+3x^{2/3}\sqrt{h}\,y+3x^{1/3}hy^2+h^{3/2}y^3)$$
$$= f(x)+(3x^{2/3}\sqrt{h}\,y+3x^{1/3}hy^2)f'(x)$$
$$+\frac{9}{2}x^{4/3}hy^2f''(x)+\theta h^{3/2}$$
$$(\theta=\theta(x,y,h) \text{ は } x \text{ をきめたとき有界})$$

であるから，上の例と同様に

$$A_sf(x) = 3x^{1/3}f'(x)+\frac{9}{2}x^{4/3}f''(x).$$

この A_s は s には無関係(これは一様性から当然の結論である)な2階微分作用素であるが，1階微分も含まれ，また係数には x が含まれる．

係数が t にも x にも関係する例としては

$$Y_t = tB_t^3$$

があり，これも Markov 過程になる．その生成作用素は

$$A_s = (x+3tx^{1/3})\frac{d}{dx}+\frac{9}{2}t^2x^{4/3}\frac{d^2}{dx^2}$$

となる．このことを証明するのは，上のように直接計算したのではかなり複雑である．後節において導入する確率微分を用いると，極めて簡単に証明される．

以上の例の一般化として，つぎの定理がなりたつ．

定理 5.33 (A. Kolmogorov) 転移確率 $p(s,x,t,E)$ が

$$\int_R p(s,x,s+h,dy)(y-x) = a(s,x)h+o(h),$$
$$\int_R p(s,x,s+h,dy)(y-x)^2 = b(s,x)h+o(h),$$
$$\int_R p(s,x,s+h,dy)(y-x)^3 = o(h)$$

$\left(\begin{array}{l}o(h) \text{ は } x \text{ には関係するが，} s \text{ について}\\ \text{は一様に } h \text{ よりも高位の無限小である}\end{array}\right)$

をみたすならば

$$A_s = a(s,x)\frac{d}{dx}+\frac{1}{2}b(s,x)\frac{d^2}{dx^2}.$$

証明
$$(p_{s,s+h}f)(x)-f(x) = \int p(s,x,s+h,dy)(f(y)-f(x))$$
$$= \int p(s,x,s+h,dy)\left[(y-x)f'(x)+\frac{1}{2}(y-x)^2f''(x)+O(y-x)^3\right]$$
$$= ha(s,x)f'(x)+\frac{1}{2}hb(s,x)f''(x)+o(h).$$

これから容易に上の式がでる．∎

上述のばあいはいずれも微分作用素になる場合であるが，差分作用素となることもある．たとえば Poisson 型の Lévy 過程 $\{X(t)\}$ は $\lambda(t)=E(X(t))$ とすると，

$$p(s,x,t,E) = p_{\lambda(t)-\lambda(s)}(E-x), \qquad p_\lambda\text{: 平均 } \lambda \text{ の Poisson 分布}$$

となるから，$\lambda(t)$ が微分可能ならば，

$$\lambda(s+h)-\lambda(s) = \lambda'(s)h+o(h),$$
$$p_{s,s+h}f(x)-f(x) = \sum_{k=0}^{\infty} e^{-(\lambda(s+h)-\lambda(s))}\frac{(\lambda(s+h)-\lambda(s))^k}{k!}[f(x+k)-f(x)]$$
$$= \lambda'(s)h[f(x+1)-f(x)]+o(h),$$
$$A_sf(x) = \lambda'(s)\Delta_1f(x), \qquad (\Delta_af)(x) = f(x+a)-f(x).$$

一般につぎの定理がなりたつことは弱収束の定義から明らかである．

定理 5.34 転移確率 $p(s,x,t,E)$ が

$$\frac{1}{h}[p(s,x,s+h,\cdot+x)-\delta_x(\cdot)] \longrightarrow \theta_{x,s}(\cdot) \qquad (h\downarrow 0)$$

$(\theta_{x,s}(\cdot)$ は $\boldsymbol{R}-\{0\}$ の上の有界測度，収束は弱収束$)$

をみたすならば

$$A_sf(x) = \int_{|u|>0}[f(x+u)-f(x)]\theta_{x,s}(du) \quad \text{すなわち} \quad A_s = \int_{|u|>0}\theta_{x,s}(du)\Delta_u.$$

Poisson 型 Lévy 過程の例は $\theta_{x,s}=\lambda'(s)\delta_1$ のばあいである．上の定理よりもっと一般な形として

$$A_sf(x) = a(s,x)f'(x)+\frac{1}{2}b(s,x)f''(x)$$
$$+\int_{|u|>0}\left[f(x+u)-f(x)-\frac{u}{1+u^2}f'(u)\right]\theta_{x,s}(du)$$
$$\left(\theta_{x,s} \text{ は測度で } \int_{|u|>0}(u^2\wedge 1)\theta_{x,s}(du)<\infty\right)$$

が考えられ，これがある意味で最も一般な生成作用素の形である．

例題 5.13 (i) 複合 Poisson 過程(λ, Φ)の生成作用素は

$$A = \lambda\int_{|u|>0}\Phi(du)\Delta_u$$

であることを示せ．

(ii) $\mathscr{F}\mu(z) = \exp\psi(z)$,

$$\psi(z) = \left[imz - \frac{v}{2}z^2 + \int_{|u|>0}\left(e^{izu} - 1 - \frac{izu}{1+u^2}\right)n(du)\right]$$

で定まる無限可解分布 μ に対応する一様 Lévy 過程 X の生成作用素 A は

$$Af(x) = mf'(x) + \frac{v}{2}f''(x) + \int_{|u|>0}\left[f(x+u) - f(x) - \frac{u}{1+u^2}f'(x)\right]n(du)$$

で与えられることを示せ．

［ヒント］ $f_z(x) = e^{izx}$ とおいて $\mu_{s,t}(E) = P\{X(t) - X(s) \in E\}$ とすれば $\mathscr{F}\mu_{s,t} = \exp[(t-s)\psi(z)]$,

$$(p_{s,s+h}f_z)(x) = \int e^{izy}\mu_{s,s+h}(dy - x) = e^{izx}\mathscr{F}\mu_{s,s+h}(z) = e^{izx}e^{h\psi(z)},$$

$$\frac{1}{h}(p_{s,s+h}f_z - f_z)(x) = e^{izx}\cdot\frac{1}{h}[e^{h\psi(z)} - 1] \longrightarrow e^{izx}\psi(z),$$

$$Af_z(x) = e^{izx}\left\{imz - \frac{v}{2}z^2 + \int_{|u|>0}\left(e^{izu} - 1 - \frac{izu}{1+u^2}\right)n(du)\right\}$$

$$= \left[m\frac{d}{dx} + \frac{v}{2}\frac{d^2}{dx^2} + \int_{|u|>0}n(du)\left(\Delta_u - \frac{u}{1+u^2}\frac{d}{dx}\right)\right]f_z(x).$$

任意の $f \in C_0^\infty(\boldsymbol{R})$ を

$$f(x) = \int e^{izx}g(z)dz = \int f_z(x)g(z)dz$$

とかいて，A を形式的に施して見よ．

(iii) Markov 過程 $X = \{X(t)\}$ のとる値は実数である必要はない．前節と本節とにのべた事柄は一般の空間（例えば $\boldsymbol{R}^n$）の値をとるばあいにも適用できる．例えば $B_1(t), B_2(t), \cdots, B_n(t)$ が独立な Wiener 過程のとき

$$B(t) = (B_1(t), B_2(t), \cdots, B_n(t))$$

を n 次元 Wiener 過程という．$\{B(t)\}$ は一様 Markov 過程で

$$p(t, x, E) = N_{0,tI}(E - x), \quad I \text{ は単位行列}$$

がその転移確率である．このとき生成作用素は

$$A = \frac{1}{2}\Delta \qquad \left(\Delta = \sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}\right)$$

で与えられることを示せ．

§5.14 確率微分方程式論の直観的背景

$\{B_1, B_2, \cdots\}$ を $[0, 1]$ の上の一様分布に従う独立な確率変数の列とする．これに対して新しい確率変数の列 $\{X_n\}$ を差分方程式

$$(\Delta) \qquad \begin{cases} X_0 = a \quad (\text{定数}), \\ X_{n+1} - X_n = f_n(X_n, B_{n+1}), \quad n = 0, 1, \cdots \end{cases}$$

で定めるとする．$f_n(X_n, B_{n+1})$ が B_{n+1} を含まなくて $f_n(X_n)$ の形をしているならば，これは通常の差分方程式で，この方程式の解 $\{X_n\}$ は通常の数列となる．このときには，方程式は，時点 n からつぎの時点 $n+1$ への変動 $X_{n+1} - X_n$ が時点 n の状態 X_n により，ある定まった法則 f_n で定められることを意味する．上の形のように $f_n(X_n, B_{n+1})$ が実際 B_{n+1} を含むならば，この方程式は，変動 $X_{n+1} - X_n$ が X_n の値以外に偶然的な影響 B_{n+1} を含むことを意味する．この意味で，差分方程式(Δ)は**確率差分方程式**とよばれる．この方程式は通常の差分方程式と同様に，n に関する帰納法によって容易に解くことができ，その解は

$$X_0 = a,$$

$$X_n = g_n(B_1, B_2, \cdots, B_n) \quad (g_n : \boldsymbol{R}^n \to \boldsymbol{R}), \quad n = 1, 2, \cdots$$

の形をしている．これは確率過程（離散時変数）$\{X_n\}$ が $\{B_n\}$ で生成される増大情報系 $\{\mathscr{F}_n(B)\}$ に従属していることを意味する．したがって $(X_1, X_2, \cdots, X_n)$ は B_{n+1} と独立である．ゆえに

$$\begin{aligned} &P\{X_{n+1} \in E \mid X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} \\ &= P\{X_n + f_n(X_n, B_{n+1}) \in E \mid X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} \\ &= P\{x_n + f_n(x_n, B_{n+1}) \in E\}. \end{aligned}$$

これから

$$\begin{aligned} P\{X_{n+1} \in E \mid X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} &= P\{X_{n+1} \in E \mid X_n = x_n\} \\ &= P\{x_n + f_n(x_n, B_{n+1}) \in E\} \end{aligned}$$

が得られるが，これは $\{X_n\}$ が

$$p(n, x, n+1, E) = P\{x+f_n(x, B_{n+1}) \in E\}, \qquad n = 1, 2, \cdots$$

を転移確率とする Markov 過程（離散時変数）となっていることを意味する．n と x とに依存する分布 $p(n, x, n+1, \cdot)$ を任意に与えたとき，上の式がなりたつような関数 $f_n(x, y)$ を求めることができる（読者試みよ．'$h_n(x, y) \equiv x+f_n(x, y)$' を y について増加関数として求め得る）から，与えられた転移確率 $p(n, x, n+1, E)$ をもつ Markov 過程を (Δ) の形の差分方程式の解として定めることができる．

以上の考察を連続時変数のばあいに適用しよう．同分布の独立確率数列 $\{B_n\}$ に相当するものは（時間的に）一様な加法過程（例えば Wiener 過程 $B=\{B_t\}$ としておこう）の時間的微小変化

$$dB_t(\omega) = B(t+dt, \omega) - B(t, \omega)$$

である．この意味はさし当り直観的に理解しておいてほしい．ただここで注意すべきは dt は離散時変数 n の増分 $(n+1)-n=1$ に相当するから，dt は正としておく．このように直観的に考えると，確率差分方程式 (Δ) は確率微分方程式

$$(\delta_0) \qquad \begin{cases} X_0 = a, \\ dX_t = f_t(X_t, dB_t) \qquad (dt>0) \end{cases}$$

となる f は任意の関数であるが，これを dB_t について Taylor 展開して

$$dX_t = f_t(X_t, dB_t) = a_1(t, X_t)dB_t + a_2(t, X_t)(dB_t)^2 + a_3(t, X_t)(dB_t)^3 + \cdots$$

が得られる．$dt \downarrow 0$ のとき $dX_t \to 0$ であるから，定数項はあらわれない．また $dB_t/\sqrt{dt}$ が Gauss 分布 $N_{0,1}$ に従うから

$$E\{|(dB_t)^n|\} = o((dt)^{n/2})$$

となり，上の展開の第3項以下は dt に対して無視してよい．したがって (δ_0) の第2式は

$$dX_t = b(t, X_t)dB_t + a(t, X_t)(dB_t)^2$$

とかいてよい．

差分方程式のときに，(Δ) を n について帰納的に解くことは，X_t を連続帰納的 $(t \to t+dt)$ に解くことで，いいかえれば確率積分方程式

$$\text{(i)} \qquad X_t = a + \int_0^t b(s, X_s)dB_s + \int_0^t a(s, X_s)(dB_s)^2$$

を解いて

$$X_t = g_t(B_s, s \leqq t)$$

の形にすることに外ならない．第1の積分は次節にのべる確率積分であるが，第2の積分は

$$V_t = \int_0^t (dB_s)^2 \qquad (t \text{ の増加関数})$$

とおくと

$$\int_0^t a(s, X_s)dV_s$$

とかける．V_t は上の定義から一見 ω に関係するが，実はつぎの補題により，$V_t = t$ となる．

補題 5.11 $$\int_0^t (dB_s)^2 = t,$$

厳密にいえば，$\Delta = \{0=t_0<t_1<\cdots<t_n=t\}$，$|\Delta| = \max_i (t_i - t_{i-1})$ とおくと

$$\operatorname*{l.i.p.}_{|\Delta|\to 0} \sum_{i=1}^n (B(t_i)-B(t_{i-1}))^2 = t.$$

証明 上の和を S_Δ とおくと，明らかに

$$E(S_\Delta) = \sum_i E(X_i^2), \qquad X_i = B(t_i) - B(t_{i-1}).$$

また X_i，$i=1, 2, \cdots, n$ は独立であるから，$\{X_i^2\}$ も独立で

$$V(S_\Delta) = \sum_i V(X_i^2) = \sum_i (E(X_i^4) - E(X_i^2)^2).$$

X_i が Gauss 分布 $N_{0, t_i - t_{i-1}}$ に従うから，

$$E(X_i) = 0, \qquad E(X_i^2) = t_i - t_{i-1}, \qquad E(X_i^4) = 3(t_i - t_{i-1})^2$$

となり，

$$E(S_\Delta) = t,$$
$$V(S_\Delta) = 2\sum_i (t_i - t_{i-1})^2 \leqq 2|\Delta| \sum_i (t_i - t_{i-1}) = 2|\Delta| t$$
$$\longrightarrow 0 \qquad (|\Delta| \to 0).$$

これから，Čebyšev の不等式により

$$\operatorname*{l.i.p.}_{|\Delta|\to 0} S_\Delta = t$$

がでる．∎

さてもとの積分方程式の第2の積分は

$$\int_0^t a(s, X_s)dV_s = \int_0^t a(s, X_s)ds$$

となるから，結局確率微分方程式(δ_0)は

$$(\delta) \qquad \begin{cases} X_0 = a, \\ dX_t = b(t, X_t)dB_t + a(t, X_t)dt, \end{cases}$$

積分形でかくと

$$(\mathrm{I}) \qquad X_t = a + \int_0^t b(s, X_s)dB_s + \int_0^t a(s, X_s)ds$$

となる．

例題 5.14 上の確率微分方程式の解が Markov 過程で，その生成作用素は

$$A_t = a(t, x)\frac{d}{dx} + \frac{1}{2}b(t, x)^2\frac{d^2}{dx^2}$$

で与えられることを直観的に説明せよ．(厳密な証明は§5.17で詳述する．)

［ヒント］ (I)を連続帰納的$(t \to t+dt)$に解くと，

$$X_t = g_t(B_s, s \leqq t)$$

となり，$(X_s, s \leqq t)$はdB_tと独立であるから

$$\begin{aligned} &P\{X_{t+dt} \in E \mid X_s = x_s, s \leqq t\} \\ &= P\{X_t + dX_t \in E \mid X_s = x_s, s \leqq t\} \\ &= P\{X_t + b(t, X_t)dB_t + a(t, X_t)dt \in E \mid X_s = x_s, s \leqq t\} \\ &= P\{x_t + b(t, x_t)dB_t + a(t, x_t)dt \in E\} \\ &= N_{m,v}(E) \qquad (m = x_t + a(t, x_t)dt, \ v = b(t, x_t)^2 dt). \end{aligned}$$

これから$\{X_t\}$は Markov 過程で，その瞬間転移確率は

$$p(t, x, t+dt, E) = N_{m,v} \qquad (m = x + a(t, x_t)dt, \ v = b(t, x)^2 dt)$$

で与えられる．これから

$$\int (y-x)p(t, x, t+dt, dy) = a(t, x)dt,$$

$$\int (y-x)^2 p(t, x, t+dt, dy) = b(t, x)^2 dt,$$

$$\int (y-x)^3 p(t, x, t+dt, dy) = 0$$

が得られる．ここで生成作用素に関する Kolmogorov の定理により，A_tが上の形にあらわされることがわかる．

§5.15 確率積分

前節で説明した確率微分方程式で

$$\int_0^t b(s, X_s)dB_s$$

の形の積分があらわれた．X_sは$\{B_u, u \leqq s\}$の関数として定まるから，$b(s, X_s)$も同様である．これは厳密ないい方をすれば，確率過程$\{b(t, X_t), t \in [0, \infty)\}$が'$B = \{B_t\}$で生成される増大情報系$\{\mathcal{F}_t(B), t \in [0, \infty)\}$'に適合するというべきであろう．したがって，

$$I_t(Y) = \int_0^t Y_s dB_s, \qquad t \in [0, \infty)$$

を$F(B) = \{\mathcal{F}_t(B)\}$に適合する$Y = \{Y_t\}$に対して定義するということに帰着される．$B$は$F(B)$に関して Wiener マルチンゲールであるということに注意すれば，もっと一般に，Bが増大情報系$F = \{\mathcal{F}_t\}$に関して Wiener マルチンゲールであるとき，Fに適合する$Y = \{Y_t\}$に対して上の積分$I_t(Y)$を定義すればよい．実際はこれだけの条件の下では$I_t(Y)$を定義することはできない．普通の積分のばあいに可測性，可積分性を仮定するように，$I_t(Y)$のばあいにも

可測性：Yは可測過程である，すなわち$Y_t(\omega)$は(t, ω)に関して可測$\mathcal{B}[0, \infty) \times \mathcal{D}(P)$である，

見本関数の局所自乗可積分性：

$$\int_0^t Y_s(\omega)^2 ds < \infty, \quad 0 < t < \infty \quad \text{a.s.}$$

を仮定する．この2条件をみたし，$F = \{\mathcal{F}_t\}$に適合する$Y = \{Y_t\}$の全体を$\mathcal{L}^2(F)$であらわす．またFに関する Wiener 過程の全体を$\mathcal{W}(F)$であらわす．かくして問題は，Fを任意の増大情報系とするとき，

$$I_t(Y) = \int_0^t Y_s dB_s, \qquad B \in \mathcal{W}(F), \qquad Y \in \mathcal{L}^2(F), \qquad t \in [0, \infty)$$

を定義することに帰着される．この定義を与え，その性質を調べることが本節の目的である．

$\mathscr{L}^2(\boldsymbol{F})$ の上で I_t を定義するにあたって，もっと狭い範囲で定義し，これを拡張していくことにする．

$\mathscr{S}(\boldsymbol{F})$ によって，つぎの形にかける $Y=\{Y_t\}$ の全体をあらわす．

“適当な $0=t_0<t_1<t_2<\cdots\to\infty$ があって

$$Y(t,\omega)=Y(t_{i-1},\omega)\in L^2(\Omega,\mathscr{F}_{t_{i-1}})\quad(t_{i-1}\leqq t<t_i),\qquad i=1,2,\cdots.\text{”}$$

$L^2(\boldsymbol{F})$ によって，つぎの条件をみたし，$\boldsymbol{F}$ に適合する可測過程の全体をあらわす．

$$E\left(\int_0^t Y_s(\omega)^2ds\right)<\infty,\qquad 0\leqq t<\infty.$$

明らかに

$$\mathscr{S}(\boldsymbol{F})\subset L^2(\boldsymbol{F})\subset\mathscr{L}^2(\boldsymbol{F})$$

である．これらの確率過程の集合は普通の線型演算に関して線型空間となっている．

$L^2(\boldsymbol{F})$ は半ノルムの列

$$\|Y\|_n=E\left(\int_0^n Y_t^2dt\right)^{1/2},\qquad n=1,2,\cdots$$

に関して完備な位相線型空間である．この位相に関して $\mathscr{S}(\boldsymbol{F})$ が $L^2(\boldsymbol{F})$ の中で稠密であることを示そう．任意の $Y\in L^2(\boldsymbol{F})$ が上の位相に関して $\mathscr{S}(\boldsymbol{F})$ の元で近似されることをいえばよい．$Y\in L^2(\boldsymbol{F})$ に対し

$$Y_t^{(m)}(\omega)=Y_t(\omega)1_{[-m,m]}(Y_t(\omega))$$

とおくと，$Y^{(m)}=\{Y_t^{(m)}\}$ は明らかに $L^2(\boldsymbol{F})$ に属し，

$$\|Y^{(m)}-Y\|_n\longrightarrow 0\qquad(\text{有界収束定理による})$$

であるから，始めから Y は有界 $(\sup\limits_{t,\omega}|Y_t(\omega)|<\infty)$ としてよい．このときには

$$Y_t^{(m)}(\omega)=m\int_{t-1/m}^t Y_s(\omega)ds$$

で定義される $Y^{(m)}$ も $L^2(\boldsymbol{F})$ に属し，すべての ω に対し，ほとんどすべての t に対し $Y_t^{(m)}(\omega)\to Y_t(\omega)$．しかも $\|Y^{(m)}\|_\infty\leqq\|Y\|_\infty$ であるから

$$\|Y^{(m)}-Y\|_n\longrightarrow 0\qquad(\text{有界収束定理による})$$

が得られる．しかも $Y^{(m)}$ は有界連続過程であるから，始めから，Y は有界連続過程としておいてよい．このような Y に対して

$$Y_t^{(m)}(\omega)=Y([tm]/m,\omega),\qquad m=1,2,3,\cdots$$

とおくと，$Y^{(m)}\in\mathscr{S}(\boldsymbol{F})$ で，しかもすべての (t,ω) に関して $Y_t^{(m)}(\omega)\to Y_t(\omega)$．有界収束定理により $\|Y^{(m)}-Y\|_n\to 0\ (m\to\infty)$ が得られる．これで $\mathscr{S}(\boldsymbol{F})$ が $L^2(\boldsymbol{F})$ の中で稠密なことが証明された．

さて確率積分

$$I_t(Y)=\int_0^t Y_sdB_s,\qquad 0\leqq t<\infty$$

の定義をしよう．まず $Y\in\mathscr{S}(\boldsymbol{F})$ に対して定義する．この Y は $\mathscr{S}(\boldsymbol{F})$ の定義でのべた形にかけるから，

$$I_t(Y)=\sum_{i=1}^{n-1}Y(t_{i-1})(B(t_i)-B(t_{i-1}))+Y(t_{n-1})(B(t)-B(t_{n-1}))$$
$$(t_{n-1}\leqq t<t_n),\qquad n=1,2,\cdots.$$

さてこの Y は $\{t_i\}$ を含む列 $\{0=t_0'<t_1'<\cdots\}$ に対しても前と同様の形にあらわされ，それに応じて $I_t(Y)$ を上のように定義しても，同じ値が得られる．すなわち，$I_t(Y)$ $(Y\in\mathscr{S}(\boldsymbol{F}))$ は Y の表現法には無関係に確定する．このことから，$I_t(Y)$ が Y に関して線型であることがわかる．また $I_t(Y)$ は $L^2(\Omega,P)$ に属し，しかも

$$\|I_t(Y)\|^2=\|Y\|_t^2\quad\left(\equiv E\left(\int_0^t Y_s^2ds\right)\right)$$

となることを示そう．

$$Y_{i-1}=Y(t_{i-1}),\qquad i=1,2,\cdots,n,$$

$$B_i=B(t_i)-B(t_{i-1}),\qquad i=1,2,\cdots,n-1,\qquad B_n=B(t)-B(t_{n-1}),$$

$$\|I_t(Y)\|^2=E\left\{\left(\sum_{i=1}^n Y_{i-1}B_i\right)^2\right\}=E\left\{\sum_i Y_{i-1}{}^2B_i{}^2+2\sum_{i<j}Y_{i-1}B_iY_{j-1}B_j\right\}.$$

Y_{i-1}, B_i は独立，$Y_{i-1}B_iY_{j-1}$ と B_j とは独立であるから，

$$E(Y_{i-1}{}^2B_i{}^2)=E(Y_{i-1}{}^2)E(B_i{}^2)=E(Y(t_{i-1})^2)(t_i-t_{i-1}),$$

明らかに $E(|B_i|)<\infty$．また $i<j$ のとき

$$\begin{aligned}&E(|Y_{i-1}B_iY_{j-1}|)\\&\quad\leqq\frac{1}{2}(E(Y_{i-1}{}^2B_i{}^2)+E(Y_{j-1}{}^2))\\&\quad=\frac{1}{2}(E(Y_{i-1}{}^2)E(B_i{}^2)+E(Y_{j-1}{}^2))<\infty,\end{aligned}$$

したがって

$$E(Y_{i-1}B_iY_{j-1}B_j)=E(Y_{i-1}B_jY_{j-1})E(B_j)=0.$$

これから

$$\|I_t(Y)\|^2=\sum_i E(Y(t_{i-1})^2)(t_i-t_{i-1})=E\left(\int_0^t Y_s^2ds\right)=\|Y\|_t^2$$

を得る．

これで t をきめたとき，$I_t:\mathcal{S}(F)\to L^2(\Omega,P)$ が線型かつ $\|I_t(Y)\|^2=\|Y\|_t^2$ であることがわかった．さてこの事実と，$\mathcal{S}(F)$ が $L^2(F)$ の中で稠密なことに着目すると，上記の I_t を $L^2(F)$ の上にまで上の性質を保ちながら拡張できる．実際 $Y\in L^2(F)$ に対して，

$$\|Y^{(m)}-Y\|_n\longrightarrow 0\quad(m\to\infty),\qquad n=1,2,\cdots$$

となる $Y^{(m)}\in\mathcal{S}(F)$ をとると，$\|Y^{(m)}-Y^{(m')}\|_n\to 0\ (m,m'\to\infty)$，$n=1,2,\cdots$ であるから，

$$\begin{aligned}\|I_t(Y^{(m)})-I_t(Y^{(m')})\|&=\|Y^{(m)}-Y^{(m')}\|_t\leq\|Y^{(m)}-Y^{(m')}\|_n\\ &\qquad(n\text{ は }t\text{ より大きい整数})\\ &\longrightarrow 0\qquad(m,m'\to\infty).\end{aligned}$$

したがって $\{I_t(Y^{(m)})\}$ は $L^2(\Omega,P)$ の中の Cauchy 列である．その極限を $I_t(Y)$ とおけばよい．この極限は近似列 $\{Y^{(m)}\}$ のとり方に無関係に Y のみで定まることは容易にわかる．

さて $Y\in\mathcal{S}(F)$ のとき，$\{I_t(Y),0\leq t<\infty\}$ がマルチンゲールであることは，上述の Y の表現から容易にわかる．これから，$Y\in L^2(F)$ のときには，上記の近似列 $\{Y^{(m)}\}\subset\mathcal{S}(F)$ をとると，

$$\|I_t(Y^{(m)})-I_t(Y)\|\longrightarrow 0$$

により，$I_t(Y)$ は可測 $\mathcal{F}_t$ である．しかも $t>s$ のとき

$$E(I_t(Y^{(m)}),A)=E(I_s(Y^{(m)}),A),\qquad A\in\mathcal{F}_s$$

がなりたつから，

$$\begin{aligned}|E(I_t(Y^{(m)}),A)-E(I_t(Y),A)|&\leq E(|I_t(Y^{(m)})-I_t(Y)|)\\ &\leq\|I_t(Y^{(m)})-I_t(Y)\|\longrightarrow 0\end{aligned}$$

を利用して，

$$E(I_t(Y),A)=E(I_s(Y),A),\qquad A\in\mathcal{F}_s$$

を得る．これで $Y\in L^2(F)$ のときにも，$\{I_t(Y),0\leq t<\infty\}$ がマルチンゲールであることがわかる．

$Y\in\mathcal{S}(F)$ のときには，上記の Y の表現から，$\{I_t(Y)\}$ が連続過程であることは明らかである．$Y\in L^2(F)$ のときには $I_t(Y)$ は $L^2(\Omega,P)$ の元として(したがって P 零測度を除いて)定められたにすぎないから，$\{I_t(Y)\}$ が連続過程であるか否かを問うことは意味がない．しかし Y の近似列 $\{Y^{(m)}\}$ のとり方を適当にとって，$\{I_t(Y^{(m)})\}$ の極限が連続過程になることを示そう．このためにつぎの補題を利用する．

補題 5.12 $\{X_t,t\in[0,a]\}$ が $E(X_t^2)<\infty$ をみたすマルチンゲールでかつ連続過程であれば

$$\varepsilon^2P\left\{\sup_{0\leq t\leq a}|X_t|>\varepsilon\right\}\leq E(X_a^2).$$

証明 $[0,a]$ の中に任意の $t_1<t_2<\cdots<t_n$ をとると，$X(t_1)^2,X(t_2)^2,\cdots,X(t_n)^2$ は劣マルチンゲールであるから，Doob の不等式により

$$\varepsilon^2P\left\{\sup_i|X(t_i)|>\varepsilon\right\}\leq E(X(t_n)^2)\leq E(X_a^2).$$

今 $\{t_1,t_2,\cdots,t_n\}\subset\boldsymbol{Q}$ のあらゆるとり方に関して左辺の上限をとれば，$\{X_t\}$ が連続過程であることから，補題の不等式の左辺が得られる．▮

さて $Y\in L^2(F)$ に対し $\{Y^{(n)}\}\subset\mathcal{S}(F)$ を適当にとると

$$\|Y^{(m)}-Y\|_m\leq\frac{4^{-m}}{2},\qquad m=1,2,\cdots$$

とすることができる．任意に固定した n に対し，$m\geq n$ である限り

$$\begin{aligned}\|Y^{(m+1)}-Y^{(m)}\|_n&\leq\|Y^{(m+1)}-Y\|_n+\|Y^{(m)}-Y\|_n\\ &\leq\|Y^{(m+1)}-Y\|_{m+1}+\|Y^{(m)}-Y\|_m<4^{-m}.\end{aligned}$$

上の補題により，$m\geq n$ である限り

$$\begin{aligned}4^{-m}P\left\{\sup_{0\leq t\leq n}|I_t(Y^{(m+1)}Y^{(m)})|\geq 2^{-m}\right\}&\leq\|I_n(Y^{(m+1)}-Y^{(m)})\|^2\\ &\leq\|Y^{(m+1)}-Y^{(m)}\|_n^2<16^{-m},\end{aligned}$$

すなわち

$$P\left\{\sup_{0\leq t\leq n}|I_t(Y^{(m+1)})-I_t(Y^{(m)})|\geq 2^{-m}\right\}\leq 4^{-m}\qquad(m\geq n).$$

Borel-Cantelli の補題により，$I_t(Y^{(m)})$, $m=1,2,\cdots$ が $t\in[0,n]$ に関して一様に収束する確率は1である．この事象を Ω_n とすれば，$\Omega_\infty=\bigcap_n \Omega_n$ の上で，'$I_t(Y^{(m)})$, $m=1,2,\cdots$ が $t\in[0,\infty)$ に関して広義一様収束する'．また明らかに $P(\Omega_\infty)=1$ である．Ω_∞ の上では $I_t(Y)=\lim_{m\to\infty} I_t(Y^{(m)})$, $\Omega_\infty{}^c$ では $I_t(Y)=0$ とおけば，$\{I_t(Y)\}$ は連続過程で，しかもおのおのの t に対し $\|I_t(Y^{(m)})-I_t(Y)\|\to 0$ となるから，$\{I_t(Y)\}$ はマルチンゲールでもある．以後 $I_t(Y)$ はこのように定義したものとすると，上の考察を綜合して，つぎの定理を得る．

定理 5.35　$I_t(Y)$, $Y\in L^2(F)$ はつぎの性質をもつ．

(i) $I_t(Y)$ は t に関して連続，Y に関して線型であり，$\{I_t(Y), t\in[0,\infty)\}$ は F に関してマルチンゲールである．

(ii) $\varepsilon^2 P\left\{\sup_{0\leqq s\leqq t}|I_t(Y)|>\varepsilon\right\}\leqq E(I_t(Y)^2)=E\left(\int_0^t Y_s{}^2 ds\right)$.

(iii) 与えられた $Y, \bar{Y}\in L^2(F)$ に対し
$$P\{(\forall t)\ \text{'}Y_s(\omega)=\bar{Y}_s(\omega), s\leqq t\text{'}\Longrightarrow \text{'}I_t(Y)=I_t(\bar{Y})\text{'}\}=1.$$

証明　(i), (ii) は上の説明に完全に含まれているから，(iii) のみを証明する．'$\alpha(\omega)\Longrightarrow\beta(\omega)$' は，いうまでもなく，'$\alpha(\omega)^{\neg}\vee\beta(\omega)$' を意味する．$I_t(Y)$ の t に関する連続性により，$(\forall t)$ は $(\forall t\in \boldsymbol{Q}\cap[0,\infty))$ でおきかえてよい．$\boldsymbol{Q}$ は可算であるから，個々の t に対し
$$P\{\text{'}Y_s(\omega)=\bar{Y}_s(\omega), s\leqq t\text{'}\Longrightarrow \text{'}I_t(Y)=I_t(\bar{Y})\text{'}\}=1$$
をいえば，十分である．

さて $\mathcal{S}(F)$ が $L^2(F)$ で稠密であることを示すときに用いた $Y\in L^2(F)$ の近似列 $Y^{(n)}\in\mathcal{S}(F)$ の求め方をよく見ると，固定した t と固定した $Y, \bar{Y}\in L^2(F)$ に対しては
$$Y_s(\omega)=\bar{Y}_s(\omega),\qquad s\leqq t$$
のときには，
$$Y_s{}^{(n)}(\omega)=\bar{Y}_s{}^{(n)}(\omega),\qquad s\leqq t$$
となるようにとることができる．$A=\{Y_s(\omega)=\bar{Y}_s(\omega), s\leqq t\}$ とおくと
$$\begin{aligned}E((I_t(Y^{(n)})-I_t(Y))^2, A)&\leqq\|I_t(Y^{(n)})-I_t(Y)\|^2\\&=\|Y^{(n)}-Y\|_t{}^2\longrightarrow 0\qquad(n\to\infty),\end{aligned}$$
同様に

$$E((I_t(\bar{Y}^{(n)})-I_t(\bar{Y}))^2, A)\longrightarrow 0.$$
A の上では $Y_s{}^{(n)}(\omega)=\bar{Y}_s{}^{(n)}(\omega)$, $s\leqq t$ で，しかも $Y^{(n)}, \bar{Y}^{(n)}\in\mathcal{S}(F)$ であるから，$\mathcal{S}(F)$ の上の I_t の定義により，上に示した $Y_s{}^{(n)}(\omega)=\bar{Y}_s{}^{(n)}(\omega)$, $s\leqq t$, $\omega\in A$ から，$I_t(Y^{(n)})=I_t(\bar{Y}^{(n)})$, $\omega\in A$ がでる．したがって上の第2式は
$$E((I_t(Y^{(n)})-I_t(\bar{Y}))^2, A)\longrightarrow 0$$
とかくこともできて，第1式と比較して $E((I_t(\bar{Y})-I_t(Y))^2, A)=0$ がでる．これは '$I_t(\bar{Y})=I_t(Y)$ が A の上のほとんどすべて (P) の ω に対してなりたつ' ことを意味する．すなわち，
$$P\{A^c\vee\{I_t(\bar{Y})=I_t(Y)\}\}=1.\qquad \blacksquare$$

さていよいよ本節の目的の $Y\in\mathcal{L}^2(F)$, $B\in\mathcal{W}(F)$ に対して，確率積分 $I_t(Y)$ を定義しよう．Y の見本関数が局所自乗可積分性をもつから，
$$Y^{(n)}(t,\omega)=Y(t,\omega)\cdot 1_{[0,n]}\left(\int_0^t Y(s,\omega)^2 ds\right),\qquad n=1,2,\cdots$$
は
$$\int_0^t Y^{(n)}(s,\omega)^2 ds\leqq n,\qquad n=1,2,\cdots$$
をみたし，$Y^{(n)}$ $(n=1,2,\cdots)$ は $L^2(F)$ に属する．したがって $I_t(Y^{(n)})$ は定義される．上の定理の (iii) により，$m>n$ のとき
$$\begin{aligned}\int_0^t Y(s,\omega)^2 ds\leqq n&\Longrightarrow Y^{(n)}(s,\omega)=Y^{(m)}(s,\omega),\ s\leqq t\\&\Longrightarrow I_t(Y^{(n)})=I_t(Y^{(m)}).\end{aligned}$$
このことから
$$I_t(Y)=I_t(Y^{(n)})\quad\left(\int_0^t Y(s,\omega)^2 ds\leqq n\text{ のとき}\right),\qquad n=1,2,\cdots$$
とおいて，$I_t(Y)$ $(Y\in\mathcal{L}^2(F),\ t\in[0,\infty))$ を矛盾なく定義することができる．

定理 5.36　確率積分 $I_t(Y)$, $Y\in\mathcal{L}^2(F)$ はつぎの性質をもつ．

(i) $I_t(Y)$ は t に関して連続，Y に関して線型であり，

(ii) 任意の $k=1,2,\cdots$ に対し
$$\int_0^k (Y_n(t))^2 dt\longrightarrow 0\ \text{i.p.}\Longrightarrow\sup_{0\leqq t\leqq k}|I_t(Y_n)|\longrightarrow 0\ \text{i.p.},$$

(iii) 与えられた $Y, \bar{Y}\in\mathcal{L}^2(F)$ に対し，

$$P\{(\forall t)\ `Y_s(\omega)=\bar{Y}_s(\omega),\ s\leqq t' \Longrightarrow `I_t(Y)=I_t(\bar{Y})'\}=1.$$

証明　(i), (iii) はそれぞれ前定理 (i), (iii) と $I_t(Y)$ ($Y\in\mathcal{L}^2(\boldsymbol{F})$) の定義から容易にでるから，(ii) のみ証明する．任意の $\varepsilon>0$ に対し

$$Y_n{}^{\varepsilon}(t,\omega)=Y_n(t,\omega)\cdot 1_{[0,\varepsilon)}\left(\int_0^t Y_n(s,\omega)^2 ds\right)$$

とおくと，仮定により，十分大きい $n_0(\varepsilon)$ に対し

$$P\{Y_n{}^{\varepsilon}(t)\neq Y_n(t),\ t\leqq k\}<\varepsilon,\qquad n>n_0(\varepsilon).$$

前定理 (ii) により

$$P\left\{\sup_{0\leqq t\leqq k}|I_t(Y_n{}^{\varepsilon})|>\varepsilon^{1/3}\right\}\leqq\varepsilon^{-2/3}E\left\{\int_0^k (Y_n{}^{\varepsilon}(t))^2dt\right\}$$
$$\leqq\varepsilon^{-2/3}\cdot\varepsilon=\varepsilon^{1/3}.$$

したがって

$$P\left\{\sup_{0\leqq t\leqq k}|I_t(Y_n)|>\varepsilon^{1/3}\right\}<\varepsilon^{1/3}+\varepsilon\qquad(n>n_0(\varepsilon)).\quad\blacksquare$$

上の定義では積分区間を $[0,t]$ としたが，同様の方法で

$$I_{s,t}(Y)=\int_s^t Y_u dB_u\qquad(s<t)$$

を定義することができる．明らかに

$$I_{s,t}(Y)=I_t(Y)-I_s(Y)$$

がなりたつ．

$$Y^{(s)}(t)=Y(s+t),\qquad B^{(s)}(t)=B(s+t)-B(s),$$
$$\boldsymbol{F}^{(s)}=\{\mathcal{F}_t{}^{(s)}=\mathcal{F}_{s+t},\ t\in[0,\infty)\}$$

とおくと，$Y\in\mathcal{L}^2(\boldsymbol{F})$，$B\in\mathcal{W}(\boldsymbol{F})$ ならば，

$$Y^{(s)}\in\mathcal{L}^2(\boldsymbol{F}^{(s)}),\qquad B^{(s)}\in\mathcal{W}(\boldsymbol{F}^{(s)}),$$
$$I_{t-s}(Y^{(s)})=I_{s,t}(Y)$$

となる．

$I_{s,t}(Y)=I_t(Y)-I_s(Y)$ により，積分区間による加法性

$$I_{s,t}(Y)=I_{s,u}(Y)+I_{u,t}(Y)\qquad(s<u<t)$$

がなりたつ．

例題 5.15　(i)　Y が $\boldsymbol{F}=\{\mathcal{F}_t\}$ に適合する連続過程であれば，$Y\in\mathcal{L}^2(\boldsymbol{F})$ で，

$$I_{s,t}(Y)=\operatorname*{l.i.p.}_{|\Delta|\to 0}\sum_{i=1}^n Y(t_{i-1})(B(t_i)-B(t_{i-1}))$$

がなりたつことを示せ．ここで $\Delta=\{s=t_0<t_1<\cdots<t_n=t\}$ は $[0,t]$ の任意の有限分割，$|\Delta|=\max_i(t_i-t_{i-1})$ とする．

［ヒント］　$\delta=\{s=t_0<t_1<\cdots\to\infty\}$ ($|\delta|=\sup_i(t_i-t_{i-1})<\infty$) に対し

$$Y^{\delta}(t,\omega)=\sum_{i=1}^{\infty}Y(t_{i-1},\omega)\cdot 1_{(-1/|\delta|,1/|\delta|)}(Y(t_{i-1},\omega))\cdot 1_{[t_{i-1},t_i)}(t)$$

とおくと，$Y^{\delta}\in\mathcal{S}(\boldsymbol{F})$ ($\subset\mathcal{L}^2(\boldsymbol{F})$) で，$|\delta|=\sup_i(t_i-t_{i-1})$ に対し，

$$Y^{\delta}(t,\omega)\longrightarrow Y(t,\omega)\qquad(|\delta|\to 0)$$

がでることは Y の連続性から明らかである．これから Y の可測性がでる．Y の見本関数の局所自乗可積分性は明らかであるから，$Y\in\mathcal{L}^2(\boldsymbol{F})$．定理 5.35 (ii) により $I_t(Y^{\delta})\to I_t(Y)$ i.p. ($|\delta|\to 0$) がでる．これから問題の等式を出すのは容易である．

注意　Y^{δ} の定義で $Y(t_{i-1},\omega)$ をとったのは Y^{δ} を $\boldsymbol{F}$ に適合させるためである．$Y(\tau_i,\omega)$ ($t_{i-1}\leqq\tau_i<t_i$) をとっても，$Y^{\delta}(t,\omega)\to Y(t,\omega)$ はでるが，Y^{δ} が $\boldsymbol{F}$ に適合するという重要な点は保証されない．

(ii)　前問と補題 5.11 を用いて，つぎの等式を示せ．

$$\int_s^t B_u dB_u=\frac{1}{2}(B_t{}^2-B_s{}^2)-\frac{1}{2}(t-s)\qquad(t>s).$$

注意　$f(t)$ が t の有界変動であれば

$$\int_s^t f_u df_u=\frac{1}{2}(f_t{}^2-f_s{}^2)$$

となるが，B_t の見本関数は有界変動でないために右辺の第2項があらわれる．

［ヒント］

$$B(t_{i-1})(B(t_i)-B(t_{i-1}))=\frac{1}{2}(B(t_i)^2-B(t_{i-1})^2)-\frac{1}{2}(B(t_i)-B(t_{i-1}))^2.$$

(iii)　$Y\in\mathcal{L}^2(\boldsymbol{F})$ のときには，$\{I_t(Y),\ t\in[0,\infty)\}$ は必ずしもマルチンゲールにならないが，$\boldsymbol{F}$ 止め時の列 $\{\tau_1\leqq\tau_2\leqq\cdots\to\infty\}$ が存在して，$\{I_{t\wedge\tau_n}(Y),\ t\in[0,\infty)\}$ はすべての n に対して $\boldsymbol{F}$ に関してマルチンゲールとなる．

［ヒント］　$\tau_n=\sup\left\{t\,\middle|\,\int_0^t Y_s^2 ds\leqq n\right\}$ とおけば $Y^{(n)}(t)=Y(t)\cdot 1_{[0,\tau_n]}(t)$ は $L^2(\boldsymbol{F})$ に属し，

$$I_{t\wedge\tau_n}(Y) = I_t(Y^{(n)}), \quad 0 \leq t < \infty \quad \text{a.s.}$$

§5.16 確率微分

増大情報系 $\boldsymbol{F}=\{\mathcal{F}_t, t\in[0,\infty)\}$ に対し，$\boldsymbol{F}$ に関する Wiener マルチンゲールの族を $\mathcal{W}(\boldsymbol{F})$ とかくことは前にのべた．本節では独立な $B^i\in\mathcal{W}(\boldsymbol{F})$，$i=1,2,\cdots,n$ を任意に固定して話をすすめる．また局所有界な見本関数をもち，$\boldsymbol{F}$ に適合する可測過程の全体を $\mathcal{B}=\mathcal{B}(\boldsymbol{F})$ であらわす．$\mathcal{B}$ は通常の（(t,ω) ごとの）加法，乗法に関し'単位をもつ可換環'である．

積分表示

(i) $$X_t = X_0+\sum_{i=1}^n\int_0^t Y_i(s)dB^i(s)+\int_0^t Y(s)ds, \quad t\in[0,\infty)$$
$$(Y_1, Y_2, \cdots, Y_n, Y\in\mathcal{B}(\boldsymbol{F}))$$

をもち，$\boldsymbol{F}$ に適合する確率過程 X の全体を $Q=Q(\boldsymbol{F})=Q(\boldsymbol{F}, B^1, B^2, \cdots, B^n)$ であらわす．$\mathcal{B}(\boldsymbol{F})\subset\mathcal{L}^2(\boldsymbol{F})$ であるから，上の確率積分は意味をもち，また最後の積分は見本関数の Lebesgue 積分の意味で定義される．（局所有界ならば当然局所可積分となることに注意せよ．）

B^i $(i=1,2,\cdots,n)$，$X(t)\equiv t$ なる確率過程 X（これを単に t とかく），上の (i) の右辺にあらわれる積分はいずれも Q に属することは明らかである．

$X\in Q$ に対し

$$(dX)(s,t) = X(t,\omega)-X(s,\omega) \qquad (t>s)$$

で定義される加法的区間関数 dX の全体を dQ であらわす．$X\neq Y$ でも $dX=dY$ となることがある．実際 $dX=dY$ となるためには，

$$Y_t(\omega) = X_t(\omega)+A(\omega), \quad 0\leq t<\infty,$$
$A(\omega)$ は可測 $\mathcal{F}_0$ な確率変数

とかけることが，必要十分である．

$X, Y\in Q$ のときには $Z_t=\alpha X_t+\beta Y_t$ (α, β: 定数) も Q に属する．このときは

$$dZ = \alpha\cdot dX+\beta\cdot dY$$

となることは定義より明らかである．したがって dQ は線型空間（ベクトル空間）である．

$X\in Q$ が積分表示 (i) をもつとき，$Z\in\mathcal{B}$ に対し

$$\bar{X}_t = \bar{X}_0+\sum_{i=1}^n\int_0^t ZY_i dB^i+\int_0^t ZYds, \quad 0\leq t<\infty$$

とおくと，$\bar{X}=\{\bar{X}_t\}$ も Q に属する．このとき

$$Z\cdot dX = d\bar{X}$$

と定義する．$Z=\alpha$（定数）のときに $Z\cdot dX$ は上述の $\alpha\cdot dX$ に外ならない．明らかに

$$Z\cdot(dX_1+dX_2) = Z\cdot dX_1+Z\cdot dX_2,$$
$$(Z_1+Z_2)\cdot dX = Z_1\cdot dX_1+Z_2\cdot dX_2,$$
$$(Z_1Z_2)\cdot dX = Z_1\cdot(Z_2\cdot dX)$$

であるから，dQ は $\mathcal{B}$ 加群（$\mathcal{B}$ の上の加群）である．

$X, \bar{X}\in Q$ が (i) の形にあらわされるとする．ただし $\bar{X}$ に対しては Y_i, Y をそれぞれ $\bar{Y}_i, \bar{Y}$ でおきかえる．このとき

$$\langle X, \bar{X}\rangle(t,\omega) = \sum_{i=1}^n\int_0^t Y_i(s)\bar{Y}_i(s)ds$$

とおくと，$\langle X, \bar{X}\rangle$ も Q に属する確率過程である．したがって $d\langle X, \bar{X}\rangle$ は dQ に属するが，これを $dX\cdot d\bar{X}$ とかくことにする．明らかに

$$dX_1\cdot dX_2 = dX_2\cdot dX_1, \qquad dX\cdot(dX_1+dX_2) = dX\cdot dX_1+dX\cdot dX_2,$$
$$(YdX_1)\cdot dX_2 = Y\cdot(dX_1\cdot dX_2), \qquad (dX_1\cdot dX_2)\cdot dX_3 = dX_1\cdot(dX_2\cdot dX_3) = 0$$

がなりたつから，$\mathcal{B}$ 加群 dQ にこの積演算をそえて考えると，dQ は $\mathcal{B}$ 多元環（$\mathcal{B}$ の上の多元環）となっていることがわかる．とくに最後の等式により，この $\mathcal{B}$ 多元環が極めて簡単な構造をもっていることがわかる．

つぎの定理は $d\langle X, \bar{X}\rangle$ を $dX\cdot d\bar{X}$ とかくことの妥当性を保証する．

定理 5.37 任意の $s<t$ に対し

$$\underset{|\Delta|\to 0}{\text{l.i.p.}}\sum_{i=1}^m(X(t_i)-X(t_{i-1}))(\bar{X}(t_i)-\bar{X}(t_{i-1})) = (dX\cdot d\bar{X})(s,t).$$

ここで Δ は $[s,t]$ の任意の分割 $\{s=t_0<t_1<\cdots<t_m=t\}$ を $|\Delta|$ は $\max_i(t_i-t_{i-1})$ をあらわす．——

この定理は補題 5.11 ($X=\bar{X}=B\in\mathcal{W}(\boldsymbol{F})$ のとき) の拡張になっている．証明は省略する．

つぎの定理は確率微分に関する最も重要な定理で普通の微積分学における微分の変換公式

$$dF(x(t)) = \sum_i \partial_i F(x(t))dx^i(t), \qquad x(t) = (x^1(t), x^2(t), \cdots, x^n(t))$$

に相当する．

定理 5.38 $F(x^1, x^2, \cdots, x^n)$ が2階連続微分可能であれば，$X^1, X^2, \cdots, X^n \in Q$ のとき，$X_t = F(X(t))$，$X(t) = (X^1(t), X^2(t), \cdots, X^n(t))$，も Q に属し，

$$dX = \sum_i \partial_i F \cdot dX^i(t) + \frac{1}{2}\sum_{i,j}\partial_i\partial_j F \cdot dX^i \cdot dX^j.$$

ここに $\partial_i F, \partial_i\partial_j F$ は $X(t)$ における値をとる．証明には F を Taylor 展開して2次の項までとり，残りの項が無視できることを示す．2次の項が無視できないのは Wiener 過程の性質 $(dB)^2 = dt$ によるのである．証明は技術的にかなり複雑であるから省略する．

上の定理では $\partial_i\partial_j F$ を含む項があらわれ普通の変換公式と異なっている．$X, Y \in Q$ のとき，

$$Y \circ dX = Y \cdot dX + \frac{1}{2}dX \cdot dY$$

とおくと，この乗法についても，dQ は Q 多元環となる．

$$(Y \circ dX)(s, t) = \lim_{|\varDelta| \to 0}\sum_{i=1}^{m}\frac{1}{2}(Y(t_{i-1}) + Y(t_i))(X(t_i) - X(t_{i-1})),$$

$$\varDelta = \{s = t_0 < t_1 < \cdots < t_m = t\}, \qquad |\varDelta| = \max_i (t_i - t_{i-1})$$

となる．右辺は Stratonovich が導入した対称確率積分に相当するものである．これを用いると次の定理が得られる．

定理 5.39 $F(x^1, x^2, \cdots, x^n)$ が無限回微分可能ならば，$X^1, X^2, \cdots, X^n \in Q$ のとき，$X_t = F(X_t^1, X_t^2, \cdots, X_t^n)$ は Q に属し，

$$dX = \sum_i \partial_i F \circ dX^i.$$

証明 右辺を変形して

$$\sum_i\left(\partial_i F \cdot dX^i + \frac{1}{2}d(\partial_i F) \cdot dX^i\right)$$

$$= \sum_i \partial_i F dX^i + \frac{1}{2}\sum_i \partial_j\partial_i F dX^i dX^j + \frac{1}{4}\sum_{i,j,k}\partial_i\partial_j\partial_k F dX^i dX^j dX^k$$

$$= \sum_i \partial_i F dX^i + \frac{1}{2}\sum_{i,j}\partial_i\partial_j F dX^i dX^j = dX \qquad \text{(前定理による)}. \quad ▮$$

例題 3.16 $B \in \mathscr{W}(F)$ とし，また $B_1, B_2, \cdots, B_n \in \mathscr{W}(F)$ は独立とする．

(i) $d(B^2) = 2BdB + dt = 2B \circ dB$ を示せ．

(ii) $dB_i \cdot dB_j = \delta_{ij}dt$，$dB \cdot dt = dt \cdot dt = 0$ を示せ．

［ヒント］ $B_i(t) = \int_0^t 1 \cdot dB_i(s)$ に注意すれば，積 $dB_i \cdot dB_j$ の定義から，すぐにでる．

(iii) $h: \boldsymbol{R}^n \to \boldsymbol{R}$ が2階連続微分可能な関数のとき

$$d[h(B_1, B_2, \cdots, B_n)] = \sum_i \partial_i h \cdot dB_i + \frac{1}{2}\triangle h dt$$

を示せ．

(iv) $X_t = e^{B(t)}$ は

$$dX_t = X_t \circ dB_t \qquad \left(\text{または } dX_t = X_t \cdot dB_t + \frac{1}{2}X_t dt\right)$$

をみたすことを示せ．

(v) $dX_t = X_t \cdot dB_t$ を解け．

［ヒント］
$$X_t \cdot dB_t = X_t \circ dB_t - \frac{1}{2}dX_t \cdot dB_t = X_t \circ dB_t - \frac{1}{2}X_t dt$$
$$= X_t \circ d\left(B_t - \frac{1}{2}t\right),$$

したがって

$$dX_t = X_t \circ d\tilde{B}_t, \qquad \tilde{B}_t = B_t - \frac{1}{2}t$$

をとけばよい．これから

$$X_t = X_0 e^{\tilde{B}_t} = X_0 e^{B_t - t/2}$$

を得ることは定理 5.39 からわかる．実際

$$\begin{aligned} d(X_t e^{-\tilde{B}_t}) &= e^{-\tilde{B}_t} \circ dX_t + X_t d(e^{-\tilde{B}_t}) \\ &= e^{-\tilde{B}_t} \circ dX_t - X_t e^{-\tilde{B}_t} \circ d\tilde{B}_t \\ &= 0 \qquad (\because\ dX_t = X_t \circ d\tilde{B}_t), \end{aligned}$$

$$X_t e^{-\tilde{B}_t} = X_0 e^{-\tilde{B}_0} = X_0.$$

§5.17 確率微分方程式

§5.14 で直観的に導入した確率微分方程式

$$(\delta)\qquad dX_t = a(t, X_t)dt+b(t, X_t)dB_t, \qquad X_0 = x \text{ (定数)}$$

または同等な確率積分方程式

$$(\mathrm{i})\qquad X_t = x+\int_0^t a(s, X_s)dt+\int_0^t b(s, X_s)dB_s$$

は前2節の説明により厳密な意味をもつことがわかった．ここで基礎におく増大情報系 $F=\{\mathcal{F}_t\}$ としては B で生成されるものをとり，上の方程式の解 X も F に適合する確率過程の中から求める．もちろん解 X が連続過程となることも確率積分，Lebesgue 積分の定義から当然のことである．したがって $a(t, x)$，$b(t, x)$ が (t, x) について連続であれば，上の積分は確定した意味をもつ．しかし連続性の条件だけでは解の一意性は保証できないことは，$b\equiv 0$ のとき（普通の意味の常微分方程式）を考えても想像できるであろう．常微分方程式論で解の存在性と一意性を保証する最も簡単な条件は Lipschitz 条件である．確率微分方程式のときにもつぎの定理がなりたつ．

定理 5.40 $a(t, x), b(t, x)$ が (t, x) に関して連続で Lipschitz 条件

$$|a(t, x)-a(t, y)|+|b(t, x)-b(t, y)| \leqq c|x-y| \qquad (c\text{: 定数})$$

をみたすならば，(δ)（すなわち (i)）は一つしかもただ一つの解 $X=\{X_t\}$ をもつ．

証明 任意の X に対して (i) の右辺の積分を $\Phi_t(X)$, $\Psi_t(X)$ とかくと，(i) は

$$X_t = x+\Phi_t(X)+\Psi_t(X), \qquad t\in[0, \infty)$$

となる．これを解くには常微分方程式論における Picard の逐次近似法の考えを用いる．すなわち

$$X_t^{(0)} \equiv x,$$
$$X_t^{(n+1)} = x+\Phi_t(X^{(n)})+\Psi_t(X^{(n)}), \qquad n = 1, 2, \cdots$$

とおくと，連続過程の列 $X^{(n)}=\{X_t^{(n)}\}$ が得られ，帰納法を用いて

$$E\{(X_t^{(n)}-X_t^{(n-1)})^2\} \leqq \frac{(c_1t)^n}{n!},$$

$$E\{(\Phi_t(X^{(n)})-\Phi_t(X^{(n-1)}))^2\} \leqq \frac{(c_2t)^n}{n!} \qquad \text{(定理 5.35 (ii) を利用する)},$$

$$E\{(\Psi_t(X^{(n)})-\Psi_t(X^{(n-1)}))^2\} \leqq \frac{(c_3t)^n}{n!}$$

（c_1, c_2, c_3 はいずれも c, T に関係して定まる定数）

が得られる．

第2式から

$$E\left\{\sup_{0\leq t\leq T}|\Phi_t(X^{(n)})-\Phi_t(X^{(n-1)})|\right\} \leqq E\left\{\int_0^T |a(s, X_s^{(n)})-a(s, X_s^{(n-1)})|ds\right\}$$
$$\leqq \left(\frac{(c_2T)^n}{n!}\right)^{1/2}.$$

右辺を n 項とする級数は収束するから，$\Phi_t(X^{(n)})$, $n=1, 2, \cdots$ がほとんど確実に $t\in[0, T]$ に関して一様収束する．定理 5.35 (ii) を用いて，第3式から

$$P\left\{\sup_{0\leq t\leq T}|\psi_t(X^{(n)})-\psi_t(X^{(n-1)})|\geqq\left(\frac{(c_3T)^n}{n!}\right)^{1/3}\right\} \leqq \left(\frac{(c_3T)^n}{n!}\right)^{1/3}.$$

Borel-Cantelli の補題により，上の式から $\Psi_t(X^{(n)})$, $n=1, 2, \cdots$ がほとんど確実に $t\in[0, T]$ に関して一様収束する．したがって

$$X_t^{(n+1)} \quad (=x+\Phi_t(X^{(n)})+\Psi(X^{(n)})), \qquad n = 1, 2, \cdots$$

もほとんど確実に $t\in[0, T]$ で一様収束する．T は任意であるから，これから $\{X_t^{(n)}\}$, $n=1, 2, \cdots$ がほとんど確実に $t\in[0, \infty)$ で広義一様収束する．この極限を $\{X_t\}$ とすれば，広義一様収束性から

$$\int_0^t (X_s^{(n)}-X_s)^2ds \longrightarrow 0 \quad (0\leqq t<\infty) \quad \text{a.s.}$$

がなりたち，定理 5.36 (ii) により

$$\sup_{0\leq t\leq T}|\Psi_t(X^{(n)})-\Psi_t(X)| \longrightarrow 0 \text{ i.p.}$$

また $\{X_t^{(n)}\}$, $n=1, 2, \cdots$ の広義一様収束性から，$n\to\infty$ のとき

$$\Phi_t(X^{(n)}) \longrightarrow \Phi_t(X) \quad \text{(広義一様収束)} \quad \text{a.s.}$$

は自明である．したがって

$$X_t^{(n+1)} = x+\Phi_t(X^{(n)})+\Psi_t(X^{(n)})$$

において，$n\to\infty$ として

$$X_t = x+\Phi_t(X)+\Psi_t(X)$$

を得る．これで解の存在が証明された．

解が二つあるとし，これを $X=\{X_t\}$, $\bar{X}=\{\bar{X}_t\}$ とする．

$$\tau_k = \tau_k(\omega) = \sup\left\{t : \sup_{0\le s\le t}|X_s(\omega)|\le k,\ \sup_{0\le s\le t}|\bar X_s(\omega)|\le k\right\}$$

とすると，$X, \bar X$ の連続性により，'$\tau_k\uparrow\infty$ a.s.' がなりたつ．（ある k で $\tau_k=\infty$ となることもある．）今 k を固定し τ_k を単に τ とかく．$X, \bar X$ が解であることを用いて

$$X_{t\wedge\tau} = x+\Phi_{t\wedge\tau}(X_{\cdot\wedge\tau})+\Psi_{t\wedge\tau}(X_{\cdot\wedge\tau}) \qquad (X_{\cdot\wedge\tau}=\{X_{t\wedge\tau},\ t\in[0,\infty)\})$$

($\bar X$ についても同様)．これから

$$E\{|\Phi_{t\wedge\tau}(X_{\cdot\wedge\tau})-\Phi_{t\wedge\tau}(\bar X_{\cdot\wedge\tau})|\}\le cE\left\{\int_0^t|X_{s\wedge\tau}-\bar X_{s\wedge\tau}|ds\right\},$$

$$E\{(\Psi_{t\wedge\tau}(X_{\cdot\wedge\tau})-\Psi_{t\wedge\tau}(\bar X_{\cdot\wedge\tau}))^2\}\le c^2E\left\{\int_0^t(X_{s\wedge\tau}-\bar X_{s\wedge\tau})^2ds\right\}$$

($\Psi_t(X_{\cdot\wedge\tau})-\Psi_t(\bar X_{\cdot\wedge\tau})$ がマルチンゲールとなることに注意)

$$\le 2kc^2E\left\{\int_0^t|X_{s\wedge\tau}-\bar X_{s\wedge\tau}|ds\right\}.$$

したがって $0\le t\le T$ なる限り

$$E\{|X_{t\wedge\tau}-\bar X_{t\wedge\tau}|\}\le(c+2kc^2)E\left\{\int_0^t|X_{s\wedge\tau}-\bar X_{s\wedge\tau}|ds\right\}$$

$$\le c_1\int_0^t E\{|X_{s\wedge\tau}-\bar X_{s\wedge\tau}|\}ds \qquad (c_1=cT+2kc^2).$$

これから

$$E\{|X_{t\wedge\tau}-\bar X_{t\wedge\tau}|\}=0,\qquad t\in[0,T]$$

が得られる．T は任意であり，$X_{t\wedge\tau}, \bar X_{t\wedge\tau}$ は t について連続であるから，

$$X_{t\wedge\tau}=\bar X_{t\wedge\tau}\quad(0\le t<\infty)\quad\text{a.s.}$$

さて $\tau=\tau_k$ とおいて $k\to\infty$ とすれば

$$X_t=\bar X_t\quad(0\le t<\infty)\quad\text{a.s.}$$

これで一意性も証明された．▌

確率微分方程式の解が一意であることから，その解は上の定理の存在証明で用いた方法で構成される．その構成法をよく見ると，その解は

$$X=f(x,B),$$
$$f:\ \boldsymbol{R}\times C[0,\infty)\longrightarrow C[0,\infty)\qquad(\text{Borel 可測}),$$
$$\xi(s)=\eta(s),\quad s\le t \Longrightarrow f(x,\xi)(t)=f(x,\eta)(t)$$

の形にかけることがわかる．これは，"近似列 $X^{(n)}$ がこの形にかけて，しかも $X^{(n)}\to X$（広義一様）a.s. である"ことから証明される．この事実を利用して，確率微分方程式 (δ) の初期条件を

'$X_0=\Xi$　　（Ξ は B と独立な実確率変数）'

でおきかえても，解 X は一意で

$$X=f(\Xi,B)$$

とかける．（このときには方程式や解を $\tilde B_t=(\Xi,B_t)$ で生成される増大情報系 F に関して理解すべきである．）さて (δ) の解 X に対して $X^{(s)}=X_{s+t}$ を考えると，$X^{(s)}$ は

$$(\delta_s)\qquad\begin{cases} dX^{(s)} = a^{(s)}(t,X^{(s)}(t))dt+b^{(s)}(t,X^{(s)}(t))dB^{(s)}(t),\\ X_0{}^{(s)} = X(s)\end{cases}$$

$$\left(\begin{array}{l} a^{(s)}(t,x)=a(s+t,x),\qquad b^{(s)}(t,x)=b(s+t,x),\\ B^{(s)}(t,\omega)=B(s+t,\omega)-B(t,\omega),\\ \text{増大情報系は上の } F=\{\mathcal{F}_t(\tilde B)\} \text{ をずらした } F^{(s)}=\\ \{\mathcal{F}_{s+t}(\tilde B)\} \text{ を用いてよい}\end{array}\right)$$

をみたし，その解は

$$X(s)=f^{(s)}(X(s),B^{(s)})$$

の形で得られる．$X^{(s)}$ は可測 $\mathcal{F}_s$，$B^{(s)}$ は $\mathcal{F}_s$ と独立であるから

$$P\{X(s+t)\in E\,|\,\mathcal{F}_s\}=P\{f^{(s)}(X(s),B^{(s)})(t)\in E\,|\,\mathcal{F}_s\}$$
$$=P\{f^{(s)}(x,B^{(s)})(t)\in E\}|_{x=X(s)}\quad\text{a.s.}$$

となる．これから解 $X=\{X(t)\}$ は Markov 性をもち，その転移確率は

$$p(s,x,t,E)=P\{f^{(s)}(x,B^{(s)})(t-s)\in E\}=P\{f^{(s)}(x,B)(t-s)\in E\}$$

($B^{(s)}$ と B とは法則同等であることに注意)

で与えられることがわかる．とくに $a(t,x), b(t,x)$ が t を含まないときには $a^{(s)}=a$，$b^{(s)}=b$ となるから，$f^{(s)}$ も s に無関係にとれて

$$p(s,x,t,E)=P\{f(x,B)(t-s)\in E\}$$

となり，時間的に一様になる．(δ) の解とすれば，X は初期値 x に関係するから $X^{(x)}$ とかくと，

$$p(t,x,E)=P\{f(x,B)(t)\in E\}=P\{X_t{}^{(x)}\in E\}.$$

$\phi\in C_0^\infty(\boldsymbol{R})$ とすれば，(δ) の解に対して

$$d\phi(X_t)=\phi'(X_t)dX_t+\frac{1}{2}\phi''(X_t)(dX_t)^2$$

$$=\phi'(X_t)(b(t,X_t)dB_t+a(t,X_t)dt)+\frac{1}{2}\phi''(X_t)b(t,X_t)^2dt$$

$$=b(t,X_t)\phi'(X_t)dB_t+(A_t\phi)(X_t)dt,$$

$$A_t=a(t,x)\frac{d}{dx}+\frac{1}{2}b(t,x)^2\frac{d^2}{dx^2},$$

$$\int p(t,x,t+h,dy)\phi(y)-\phi(x)$$

$$=E(\phi(X_{t+h})-\phi(X_t)\,|\,X_t=x)$$

$$=E\left(\int_t^{t+h}b(s,X_s)\phi'(X_s)dB_s+\int_t^{t+h}(A_s\phi)(X_s)ds\,\middle|\,X_t=x\right)$$

$$=E\left(\int_t^{t+h}(A_s\phi)(X_s)ds\,\middle|\,\mathcal{F}_t\right)=E\left(\int_t^{t+h}(A_s\phi)(X_s)ds\,\middle|\,X_t=x\right),$$

$$\frac{1}{h}\left[\int p(t,x,t+h,dy)\phi(y)-\phi(x)\right]\longrightarrow A_t\phi(x)\qquad(h\downarrow 0).$$

したがって，(δ) の解は（初期条件が $X_0=\Xi$ のときも）Markov 過程でその生成作用素は A_t である（例題 5.14 参照）．

例題 5.17 (i) **Langevin の方程式**

$$dX_t=-cX_tdt+dB_t,\qquad X_0=\Xi\qquad(c>0)$$

を解け．

［ヒント］ $(dt)^2=0$, $dt\cdot dB_t=0$（したがって $(dX_t)^2=0$）を念頭において，上の方程式の解 X_t は

$$d[e^{ct}X_t]=cX_te^{ct}dt+e^{ct}dX_t=e^{ct}dB_t,$$

$$e^{ct}X_t=\Xi+\int_0^t e^{cs}dB_s,$$

$$X_t=e^{-ct}\Xi+\int_0^t e^{c(s-t)}dB_s.$$

(ii) $X_t=B_t^3$ のみたす確率微分方程式と $\{X_t\}$ の生成作用素とを求めよ．

［ヒント］ $dX_t=3B_t^2dB_t+3B_t(dB_t)^2=3B_t^2dB_t+3B_tdt$, したがって

$$\begin{cases} dX_t=3X_t^{2/3}dB_t+3X_t^{1/3}dt,\\ X_0=0\end{cases}$$

が求める方程式である．

注意 $X_t\equiv 0$ もこの方程式をみたすから，解は一意でない．実際 $b(t,x)=3x^{2/3}$, $a(t,x)=3x^{1/3}$ は $x=0$ で Lipschitz 条件をみたしていない．

$$\tau_a=\inf\{t\geqq a\,|\,B_t=0\}\qquad(a>0)$$

とすると

$$X_t=\begin{cases}B_t, & t\leqq\tau_a,\\ 0, & t>\tau_a\end{cases}$$

も上の方程式の解である．したがって無限個の解がある．

§5.18 1次元拡散過程

$C=C[0,\infty)$（$[0,\infty)$ 上の連続関数の全体）の上に $x\in\boldsymbol{R}$ に依存する確率測度の系 $\{P_x, x\in\boldsymbol{R}\}$ が与えられているとする．$\omega\in C$ に対して $\omega(t)$ を対応させる写像を X_t であらわす．（これは射影写像 π_t と同じであるが，ここでは X_t とかくことにする．） $\{X_t(\omega), t\in[0,\infty)\}$ は，すべての $x\in\boldsymbol{R}$ に対して，(C,P_x) の上で連続過程となっている．さて

$$\mathcal{B}_t=\sigma[X_s, s\leqq t]$$

とする．もし $P_x\{X_0=x\}=1$ と，

(M) (**Markov 性**)

$$P_x\{X_{s+t}\in E\,|\,\mathcal{B}_t\}=P_x\{X_t\in E\}|_{x=X(s)},\qquad x\in\boldsymbol{R},\quad s,t\geqq 0,\quad E\in\mathcal{B}^1$$

（右辺を $P_{X_s}\{X_t\in E\}$ とかくことがある）

を仮定すれば，すべての $x\in\boldsymbol{R}$ に対して

$$\mathcal{M}_x=\{X_t(\omega), t\in[0,\infty), \omega\in(C,P_x)\}$$

は（時間的に）一様な連続 Markov 過程となっている．しかも

$$p(t,x,E)=P_x(X_t\in E),\qquad t\in[0,\infty),\quad x\in\boldsymbol{R},\quad E\in\mathcal{B}^1$$

とおけば，これはすべての $\mathcal{M}_x$ にとって転移確率となっている．

今任意の一様な連続 Markov 過程 $Y=\{Y_t(\omega), t\in[0,\infty)\}$ がある確率空間 (Ω,P) の上に与えられ，上の $p(t,x,E)$ がその転移確率となっているならば，$Y=\{Y_t\}$ は

$$\mathcal{M}_\mu=\{X_t(\omega), t\in[0,\infty), \omega\in(C,P_\mu)\}$$

$$\left(\mu(E)=P\{X_0\in E\}\ (X\text{ の初期分布}),\quad P_\mu(\Lambda)=\int_R\mu(dx)P_x(\Lambda)\right)$$

と法則同等である．したがって $p(t,x,E)$ を転移確率とする一様な連続 Markov 過程を研究するには，上の系

$$\mathcal{M}=\{\mathcal{M}_x,\ x\in \boldsymbol{R}\}$$

の性質を研究すれば十分である．

§5.4 でのべたように $\boldsymbol{B}=\{\mathcal{B}_t\}$ に関して $\tau:\Omega\to[0,\infty]$ が停止時であるとは，

$$\{\tau\leqq t\}\in\mathcal{B}_t,\qquad t\in[0,\infty)$$

(これは $\{\tau<t\}\in\mathcal{B}_t$ と同等である)

がなりたつことである．また τ が $\boldsymbol{B}$ 停止時($\boldsymbol{B}$ に関する停止時)のとき

$$\mathcal{B}_\tau=\{B\,|\,B\cap\{\tau\leqq t\}\in\mathcal{B}_t\}$$

と定義する．さて (M) のかわりに，より強い条件

(SM) **(強 Markov 性)**

$$P_x\{X_{\tau+t}\in E\,|\,\mathcal{B}_\tau\}=P_x\{X_t\in E\}\,|_{x=X_\tau}\quad(=P_{X_\tau}\{X_t\in E\}\ \text{とかく}),$$

$x\in\boldsymbol{R}$，τ は有限な $\boldsymbol{B}$ 停止時，$E\in\mathcal{B}^1$

がなりたつとき，$\mathcal{M}=\{\mathcal{M}_x,\ x\in\boldsymbol{R}\}$ を一様な連続強 Markov 過程(または拡散過程)という．以後 $\mathcal{M}=\{\mathcal{M}_x\}$ は拡散過程とする．

$$P_x\{(\exists t)\ \ X_t>x\}>0,\qquad P_x\{(\exists t)\ \ X_t<x\}>0$$

のとき，x を $\mathcal{M}$ の正則点(regular point)という．すべての $x\in\boldsymbol{R}$ が $\mathcal{M}$ の正則点であるとき，$\mathcal{M}$ を**正則拡散過程**という．

正則拡散過程に関する Feller の理論を概説しよう．

$$\tau_a=\inf\{t>0\,|\,X_t=a\}$$

を a への到達時という．これは $\boldsymbol{B}$ 停止時である．$x\in[a,b]$ の関数

$$s(x:a,b)=P_x\{\tau_a>\tau_b\}$$

は，x が a から b へ動くとき，0 から 1 へ連続的に強義増大する．$\mathcal{M}$ の強 Markov 性を用いて，

'$x\in[a,b]\subset[a',b']$ のとき

$$s(x:a',b')=s(x:a,b)s(b:a',b')+(1-s(x:a,b))s(a:a',b')'$$

が得られる．この式は $x\in[a,b]$ では $s(x:a',b')$ が $s(x:a,b)$ の1次式であることを示している．この事実を念頭において，連続，強義増大関数 $s(x)$ を見出して，任意の $a<b$ に対し

$$s(x:a,b)=s(x)\ \text{の 1 次式}$$

とすることができる．$s(x:a,b)$ は $x=a,b$ でそれぞれ 0, 1 であるから

$$s(x:a,b)=\frac{s(x)-s(a)}{s(b)-s(a)}.$$

このような $s(x)$ があれば，$\alpha s(x)+\beta\ (\alpha>0)$ もこの性質をもつから，$s(x)$ は一意ではない．しかし，二つあれば，一方は他方の1次式である．

つぎに

$$q(x:a,b)=E_x(\tau_a\wedge\tau_b),\qquad a<x<b$$

$$\left(E_x(\varphi)=\int_\Omega\varphi(\omega)P_x(d\omega)\right)$$

とおくと，

'$x\in[a,b]\subset[a',b']$ に対し

$$\begin{aligned}q(x&:a',b')\\&=q(x:a,b)+s(x:a,b)q(b:a',b')+(1-s(x:a,b))q(a,a',b')\\&=q(x:a,b)+\frac{s(x)-s(a)}{s(b)-s(a)}q(b:a',b')+\frac{s(b)-s(x)}{s(b)-s(a)}q(a,a',b')'\end{aligned}$$

がなりたつことも強 Markov 性からでる．$q(x:a,b)>0$ であるから，上の式は $q(x:a',b')$ を x の関数(したがって $s(x)$ の関数)と見て，$s(x)$ の狭義凹関数(下から見て凹)となっている．

$q(x)\equiv q(x:a,b)$，$x\in[a,b]$ が $s(x)$ の凹関数であることがわかったから，

$$m(x)=-D_sq(x)=-\lim_{h\downarrow 0}\frac{q(x+h)-q(x)}{s(x+h)-s(x)}$$

は右連続狭義増加関数であるが，連続とは限らない．$m(x)$ で定まる Lebesgue-Stieltjes 測度を dm とする．dm は開集合に対しては正である．$m(x)$ の飛躍点では dm は正となる．$u(x)$ を有界変動な関数とする．du が dm に絶対連続なとき，その Radon-Nikodym 微分係数 du/dm を D_mu であらわす．これは dm 測度 0 を除いて定まる．

さて A を転移測度 $p(t,x,E)$ の生成作用素とする．この意味は前にもいったように

$$Af(x)=\lim_{t\downarrow 0}\frac{p_tf(x)-f(x)}{t},\qquad p_tf(x)=\int_R p(t,x,dy)f(y)$$

であるが，正確にいうには A の定義域 $\mathcal{D}(A)$ を指定する必要がある．ここでは一応 $\mathcal{D}(A)$ は，上の極限が x の連続関数となり，しかも $\pm\infty$ でそれぞれ有限極限値をもつものの全体としておく．このときには，

'$f\in\mathcal{D}(A)$ のときには，$(D_mD_sf)(x)$ が連続関数となるようにとれて，

$$(Af)(x)=(D_mD_sf)(x)$$

となる．

これが拡散過程の生成作用素の Feller の表現である．

Wiener 過程に対する転移作用素 $p(t,x,E)=N_{0,t}(E-x)$ に対しては

$$s(x)=x,\quad m(x)=2x$$

にとれて

$$A=D_mD_s=\frac{1}{2}\frac{d^2}{dx^2}$$

となる．

また確率微分方程式

$$dX_t=a(X_t)dt+b(X_t)dB_t,\quad X_0=x$$

$$(b(x)>0,\ \|a'\|_\infty, \|b'\|_\infty<\infty\ \text{とする})$$

をといて，$X=\{X_t\}$ の見本空間 C の上に $X=\{X_t\}$ の確率法則 P_x をいれると，拡散過程が得られ，$f\in C_b{}^2$ (有界な2回連続可微分関数の族)に対しては

$$Af(x)=a(x)\frac{df}{dx}+\frac{1}{2}b(x)^2\frac{d^2f}{dx^2}$$

となる．これを Feller の標準形にするには

$$\begin{aligned}
af'+\frac{1}{2}b^2f'' &= \frac{1}{2}b^2\left(f''+\frac{2a}{b^2}f'\right)\\
&=\frac{1}{2}b^2(f''+\phi'f')\qquad \left(\phi=\int\frac{2a}{b^2}dx\right)\\
&=\frac{1}{2}b^2e^{-\phi}(e^\phi f''+e^\phi\phi'f')\\
&=\frac{1}{2}b^2e^{-\phi}(e^\phi f')'\\
&=\frac{d}{2b^{-2}e^\phi dx}\left(\frac{df}{e^{-\phi}dx}\right)
\end{aligned}$$

$$=D_mD_sf\qquad \left(s=\int e^{-\phi}dx,\ \ m=\int\frac{2e^\phi}{b^2}dx\right).$$

例題 5.18　(i)　$$A=-(x+1)\frac{d}{dx}+\frac{1}{2}\frac{d^2}{dx^2}$$

の Feller 標準形を(形式的に)求めよ．またこれを生成作用素とする拡散過程を確率微分方程式を用いて定めよ．

［ヒント］　確率微分方程式の解の形から，§5.8(W.7)を用いて強 Markov 性を示せ．

(ii)　C の上の確率測度 P_x を

$$P_x(\Lambda)=P\{B_\cdot^{(x)}\in\Lambda\}$$

(B は Wiener 過程，$B_t{}^{(x)}=x+B_t$，$B_\cdot^{(x)}$ は $B^{(x)}$ の見本過程)

とすると，

$$\mathcal{M}_x=\{B_t{}^{(x)}(\omega),\ t\in[0,\infty),\ \omega\in(\Omega,P_x)\},\qquad x\in\boldsymbol{R}$$

は $p(t,x,E)=N_{0,t}(E-x)$ を推移確率とする拡散過程であることを示せ．

［ヒント］　§5.8(W.7)を用いて強 Markov 性を証明せよ．

あとがき

確率論に関する文献はここ20年位の間に急激に増加し，その全貌を明らかにすることは，望むべくもないので，本書ではその中の一部を選んで紹介した．進んで研究される方のために章を追って参考文献をあげておこう．

第1章の内容は本質的には高校数学の程度であるが，これを現代的な確率論の立場から眺めたものである．

第2章では，測度論の中で確率論ですぐ必要になる事項を説明した．

第3章では有限試行のばあいに第1章で説明したことが，一般の試行のばあいにどうなるかを示した．ここで確率測度を完全可分なものに制限するという本書の立場を説明した．応用上はこれで十分であり，条件付確率測度を見本空間の上に考えられるという点では都合がよいし，その他にもいろいろ利点はあるが，時には無理な点があることも後で気付いた．やはり制限するにしても，Kolmogorovのように完全性のみを要求する方がよかったかも知れない．条件付確率を，σ 加法族に対して定義するのは Doob の立場で，分割に対して定義するのは Kolmogorov の立場である．後者の方が精密であり，また直観ともよりよく合致しているが，問題によっては σ 加法族による方が推論形式が簡単になることも多い．結局両者を自由に使えるのが，望ましいであろう．

第4章の独立確率変数の和は現在では古典に属する事項である．本書では実確率変数のばあいに限り，理論の一部を紹介した．この分野の参考書としては

B. V. Gnedenko - A. N. Kolmogorov: Limit distributions for sums of independent random variables, Addison-Wesley, 1954,

W. Feller: An introduction to probability theory and its applications vol. 2, John Wiley, 1966

をおすすめする．

第5章は確率過程論である．この理論は現在では確率論の大部分を占めているので，全貌を紹介することは到底できない．しかし基本的なことに紙数をとりすぎて，重要なことを割愛せざるを得なかったのは残念である．§5.1から§5.4

までは後で用いるためにいろいろ基本事項について詳しく説明したが，残りの紙数もなくなって，十分でなかったように思う．§5.5, §5.6のマルチンゲールの理論はほとんどすべて J. L. Doob に負うといっても過言ではない．ここでは Doob の理論を紹介した．この方面では P. Meyer を中心とするフランス学派の研究が著しい．たとえば

P. Meyer: Probabilités et potentiel, Hermann, 1966

は Doob 以後の多くの重要な結果を含んでいる．

§5.7の Gauss 系は無限次元の Gauss 分布の理論といってもよいであろう．§5.8では Wiener 過程について，古典的な結果をのべた．Wiener 過程は確率過程の中で最も基本的なものであり，後節にのべる確率微分方程式の理論の基盤となっている．Wiener 過程に関するすぐれた参考書としては

飛田武幸：ブラウン運動，岩波書店，1975

をおすすめする．これは極めて広汎な内容を含んでいる．

§5.9の多項配置と Poisson 配置は応用上よく用いられているが，本書のように詳細に説明したものはないように思われる．

§5.10の加法過程では，P. Lévy の名著

Theorie de l'addition des variables aléatoires, Gauthier-Villars, 1937

にある Lévy の分解定理の厳密な定式化と証明を与えるため多大の頁数を要した．第4章でのべた独立確率変数の理論に関する1930年代の多くの研究は，(連続時変数の) 加法過程を念頭において，離散時変数のばあいの理論を構成していたようであるが，Lévy は直接連続時変数のものに立ち向って，透徹した理論を構築した．彼の叙述は，極めて直観的であって，定式化も証明も必ずしも明確ではなかった．その後 Kolmogorov や Doob によって確率過程論を築くのに都合のよい測度論的基盤が与えられたので，その線に沿って Lévy の思想を説明することにした．上に引用した Lévy の本には，この定理の応用として，多数の興味ある結果が出されているが，本書ではその一部を紹介したにすぎない．

§5.12, §5.13および§5.18は Markov 過程に関する序論である．例えば§5.18の拡散過程については

K. Itô-H. P. McKean: Diffusion processes and their sample paths, Springer, 1969

を参照されたい．また

E. B. Dynkin: Markov processes, I, II, Academic Press, 1965

は，それまでの Markov 過程論の成果の集大成といえよう．

§5.14から§5.16では Wiener 過程を基礎にした確率微分方程式の古典理論を幾分現代的に整理してのべた．応用上はこの程度で十分であるが，この理論をより透明かつ自然にするには，Doob が示唆したように，Wiener 過程のかわりにマルチンゲールを用いて，確率積分を定義しなければならない．これによって確率微分も極めて自然に定義される．この新しい理論と応用を組織的にのべた良書として，

渡辺信三：確率微分方程式，産業図書，1975

を推薦したい．

なおエルゴード理論，Markov 過程とポテンシャル論との関係，Markov 統計力学については触れなかったが，これらもまた極めて重要であることはいうまでもない．

末筆ながら，校正の段階で多数の誤りを訂正して下さった，名古屋大学佐藤健一教授に深甚な謝意を表したい．

欧 文 索 引

和 文 索 引

ア 行

カ 行

サ 行

タ 行

ナ 行

ハ 行

マ 行

ヤ 行

ラ 行

確率論　岩波基礎数学選書

1991年5月30日　第1刷発行 ©
2009年4月3日　第9刷発行

著　者　伊藤　清（いとう　きよし）

発行者　山口昭男

発行所　株式会社　岩波書店
〒101-8002 東京都千代田区一ツ橋 2-5-5
電話案内 03-5210-4000
http://www.iwanami.co.jp/

印刷・精興社　製本・三水舎

ISBN 4-00-007816-X　　Printed in Japan